(2)

QUADRATIC HADAMARD MEMORIES II

Hendricus G. Loos

Laguna Research Laboratory
3015 Rainbow Glen, Fallbrook CA 92028–9765

July 1990

Sponsored by
Defense Advanced Research Projects Agency (DoD)
Aerospace Technology Office
"Adaptive Stochastic Content–Addressable Memory"
ARPA Order No. 6429
Issued by Army Missile Command Under
Contract # DAAH01–88–C–0887

Technical Report #2

Approved for public release; distribution unlimited

Name of Contractor: Laguna Research Laboratory
Business Address: 3015 Rainbow Glen, Fallbrook CA 92028–9765
Effective Date of Contract: 26 Sep 1988
Contract Expiration Date: 31 Dec 1990
Principal Investigator: Hendricus G. Loos, Tel. (619) 728–8767
Work Title: "Adaptive Stochastic Content–Addressable Memory"

ABSTRACT

This report is the second part of an investigation of the so called "Dominant Label Selector", (DLS) which is the rear stage of a novel associative memory called "Selective Reflexive Memory" (SRM). The front stage of the SRM consists of a Bidirectional Linear Transformer (BLT) the output of which is processed by the DLS. The BLT transforms a bipolar input **x** into a linear combination of Hadamard vectors, and the task of the DLS is to select the dominant Hadamard vector from this linear combination. This vector is then returned to the BLT for a back stroke, which produces the stored vector closest to **x**. An attractive choice of DLS is the "Quadratic Hadamard Memory", which employs quadratic activations, and stores Hadamard vectors. Previously, this DLS was investigated by means of the asynchronous discrete model. In the ~~present~~ report the investigation of the Quadratic Hadamard Memory is extended to the continuous model in which input capacitance and resistance of amplifiers is accounted for, and the coupling between BLT and DLS can be studied.

A Liapunov function ("energy") is constructed, and it follows that the DLS is stable. Sufficient conditions for instability of stationary states are derived from the energy and also from the equations of motion, in terms of the divergence of the flow in activation space. The energy landscape is explored for the case of maximum symmetry, i.e., for zero thresholds. We find a small central crater with an undulated ridge. Gullies run in the radial direction, over the ridge, and down the outer slopes, toward the Hadamard points. The deepest gullies are those directed towards a Hadamard point. The stationary points on the ridge are unstable and found to have principal Hadamard spectra, i.e., their signals are proportional to the sum of m Hadamard vectors. For m=1, the signal is proportional to a single Hadamard vector (the spectrum is "pure"). For this case, and also for principal spectra with m=2, the signal path is a radial line. For principal spectra with larger m, the path curves in the region where the neuron output function is nonlinear. The dynamics is decomposed into longitudinal and transverse parts. This decomposition leads to an adiabatic fake dynamics, in which the signal is constraint on a hypersphere H_R, and the longitudinal dynamics is omitted. We let the signal find its transverse equilibrium on H_R before going to the next hypersphere H_{R+dR}. The succession of transverse equilibria forms the transverse adiabatic path. This path is found to link the stationary points of the true dynamics with signal points that have principal spectra in the proportional region, i.e., in the region where the signals are proportional to the activations, either exactly or approximately. It is found that for thresholds with principal spectra, the signal spectrum is conserved in the proportional region. If the BLT output **u** is applied to the DLS as external coupling, and a certain large uniform threshold term is added, then the DLS has as only bipolar stationary points the Hadamard points. However, the large uniform threshold term spoils the early dynamics, by pushing the signal point out of the gully that runs to the Hadamard point that is dominant in the BLT output **u**. To avoid this from happening, the large uniform threshold term is omitted, but then, spurious stable states are let back in. It is shown however that such spurious states are dynamically inaccessible if the external coupling constant is chosen properly, and the gain is large enough. This is shown in a tedious analysis which circumvents the need to integrate the N coupled nonlinear differential equations, something I cannot do. In the proportional region these equations can be integrated in spite of the nonlinear selfcoupling (i.e., the quadratic activation). For small coupling constant, the signals in this region undergo spectral purification, which can be made as large as desired by choosing the gain large enough. In this purification the dominant Hadamard component in the signal becomes even more dominant as time goes on. After the leaving the proportional region, a final purification takes place which makes the spectrum pure, i.e., the signal becomes a single Hadamard vector. Thus we have a proof

that a DLS of dimension N (needs to be a power of 2, but is otherwise arbitrary), externally coupled to a BLT, will provide perfect associative recall of N stored vectors, if the coupling constant is chosen properly, and the gain is large enough. Large gains are desirable, but the coupling constants specified in the theorem are much too small for practical applications.

Numerical computations have been performed for SRMs of dimension N=8 and 16, for large gains and practical values of the coupling constant. Perfect associative recall was found of N random stored bipolar vectors, for any bipolar input with a unique nearest stored vector.

Because of the length of the paper, we show an overview of the sections and the theorems, stated in abbreviated form, for the purpose of orientation only.

INTRODUCTION
SELECTIVE REFLEXIVE MEMORY (SRM)
DOMINANT LABEL SELECTOR (DLS)
CONTINUOUS MODEL FOR DLS
NEURON OUTPUT FUNCTIONS
THE ENERGY
Theorem 1: The DLS is stable.

SPECIFIC FORMS OF EQUATIONS OF MOTION AND ENERGY
COUPLING SCHEMES
DYNAMIC REGIONS
STATIONARY POINTS OF THE I_N

Theorem 2: For thresholds $r_a=N^2-4N$ + external coupling, the only stationary points of the I_N are the Hadamard points.

STABILITY
Theorem 3: Any stable point of the DLS must either lie in the region $y_1 \leq 1/2$ or in the corners of the solid hypercube.

Theorem 4: For a DLS with threshold bound of $N^2/4$ there is a stationary point in every Hadamard corner.
Theorem 5: The stationary point of Theorem 4 is stable.

ENERGY LANDSCAPE
Theorem 6: For zero thresholds, the origin in signal space is stable.
Theorem 7: For zero thresholds, the stationary states in the proportional region, and away from the origin, are unstable.
Theorem 8: The ridge set has principal spectra.
Theorem 9: For zero thresholds, the stationary points of the energy function constraint to a hypersphere, in the proportional region, have ±principal spectra.

DECOMPOSITION INTO LONGITUDINAL AND TRANSVERSE DYNAMICS
ADIABATIC FAKE DYNAMICS
Theorem 10: Through every stationary point of the true DLS dynamics goes a transverse adiabatic path, along which, in the proportional region, the transverse force vanishes.
Theorem 11: For zero threshold, the transverse equilibria in the proportional region have ±principal spectra.

INTRODUCTION

The main problems of concern in associative memories are early saturation, fault sensitivity, and hardware implementation. Hopfield memories [1] are robust, but suffer from early saturation. The latter problem is solved by ART [2], but at the cost of fault sensitivity of the upper layer. Early saturation can also be circumvented by employing vectors with dilute information [3], but this approach is wasteful of memory dimension. The use of connection matrices that are more sophisticated than the Hopfield matrix [1] also may solve this problem [4,5], but at a loss of locality of the learning rule, with unacceptable consequences for hardware implementations in applications with large dimension. The use of Coulomb like activations [6] makes it possible to load up associative memories to great density, but it forces the individual neurons to be rather complicated, with undesirable consequences for hardware implementations. Making memories bidirectional [7] does not give relief of early saturation [8].

In Phase I of the present DARPA SBIR project we outlined a new approach to associative memories which appears to have the potential of overcoming the early saturation problem, while retaining fault tolerance. The approach involves a two–stage memory, shown schematically in Fig. 1.

In Phase II of the project, a promising architecture for the rear stage was identified and investigated. The device is an associative memory with Hadamard vectors as stored states, and with the activation taken as a quadratic function of the incoming neuron signals, instead of the customary linear function. Using a quadratic activation of course carries a penalty in hardware implementation, and it may lead to proliferation of the number of connections. The latter is not found to be a problem, as the number of connections required is about the same as for a fully connected Hopfield memory. The hardware complication due to quadratic activations appears to be rather mild as compared to that due to Coulomb like activations.

The resulting "Quadratic Hadamard Memory" was investigated with the

asynchronous discrete model, as discussed in an earlier report [9]. In the present report the investigation is extended to the continuous model.

The associative memory considered (see Fig. 1) consists of a bidirectional front stage which is capable of a forward stroke and a backstroke, both of which are linear transformations. We call this frontstage a "Bidirectional Linear Transformer" (BLT). In its forward stroke, the BLT transforms the N dimensional bipolar input vector **x** into a vector **u** with integer components in the range [–N,N]. The BLT is arranged such that its rear output **u** is a linear combination of orthonormal labels of the stored states. The vector **u** is presented to the rear stage, called "Dominant Label Selector" (DLS). This device is to select the dominant label from the linear combination **u**. The labels are here chosen as Hadamard vectors. The settled DLS output **y** is to be the Hadamard vector closest to **u**. The vector **y** is returned to the BLT and processed in a backstroke, which produces from the label **y** the stored state to which it belongs. If everything works as expected, that stored state is the one closest to the input **x**. The whole device, BLT plus DLS, is called "Selective Reflexive Memory" (SRM), where "selective" indicates the selection of the dominant Hadamard vector by the DLS, and "reflexive" alludes to the bidirectional nature of the BLT. The BLT may be seen as a BAM [7] without rear thresholding. Front thresholding is optional. The SRM may be likened to an ART, in which the winner–take–all circuit in the top layer is replaced by a DLS. Since the DLS is a distributed winner–take–all, its use is expected to overcome the fault sensitivity of the ART top layer.

Conventions and notations are the much the same as in [9]. The input and output of a neuron threshold function are respectively called "activation" and "signal" of the neuron. The summation convention of tensor calculus has been used where convenient. In order to distinguish from unsummed expressions, we have used the convention in its strict form [10]: in a product, summation over a repeated index is implied only if the index appears twice, once as a subscript, and once as a superscript. For instance, $u^a v_a$ is summed, but $u_a v_a$ is not.

Indices are used as follows. i, j, and k denote components in input space. a, b, c, d, and p denote components in the space between the BLT and DLS, and also components of the DLS state vectors. α, β, and γ are used to name stored vectors and their labels, the Hadamard vectors. All indices range from 1 to N. A statement involving unspecified "life" indices [10] is meant to be true for all values 1 to N for such indices.

The Kronecker delta is written as δ with two indices. If the indices have the same value, the symbol stands for unity, else it stands for zero.

Indices are raised and lowered with the Kronecker delta as metric tensor. Hence, v^a and v_a have the same numerical value.

As a further simplification of appearance, 1 is often written as + and –1 as –, when no confusion with composition symbols can arise.

Customary mathematical shorthand is used where convenient: ϵ means "is an element of", $\forall$ means "for all", $\exists$ means "there exists", ==> means "implies", and <== means "is implied by".

In the continuous model of the DLS, the signals **y** lie either in the closed solid hypercube $J_N = [-1,1]^N$, or in the open solid hypercube $J_N' = (-1,1)^N$, depending on whether or not the sigmoidal neuron output function s(v) attains the values ±1. In the

former case, the signals **y** can settle at one of the *corner points* of the hypercube, defined as the set $I_N = \{-1,1\}^N$. The I_N is the set of signals considered in the discrete model.

SELECTIVE REFLEXIVE MEMORY

The *Selective Reflexive Memory* (SRM) consists of two stages. The front stage is a *Bidirectional Linear Transformer* (BLT) which in its forward stroke performs on the bipolar input vector **x** the linear transformation

$$u_b = B_b{}^i x_i \quad , \tag{1}$$

The connection matrix of the BLT is chosen as

$$B_b{}^i = h_b{}^\alpha q_\alpha{}^i \quad , \tag{2}$$

where $\mathbf{q}_\alpha$, $\alpha = 1$ to N, are the stored bipolar vectors, and the $\mathbf{h}_\alpha$ are Hadamard vectors. We have chosen the same dimension N for the BLT front and rear vector spaces, and have taken the number of stored states equal to N as well. It will be clear from the theory how to modify these choices if desired. The structure (2) of the BLT connection matrix is Hebbian, i.e., it can be built up adaptively by Hebb learning.

The Hadamard vectors $\mathbf{h}_\alpha$ are rows of a Hadamard matrix, i.e., an orthogonal matrix (up to a scalar factor) with entries + and −. Properties of Hadamard vectors used in this report are shown in Appendix A. The Hadamard vector $\mathbf{h}_\alpha$ serves as a label for the stored state $\mathbf{q}_\alpha$.

With the connection matrix (2), Eq. (1) gives for the rear output of the BLT

$$u_b = h_b{}^\alpha q_\alpha{}^i x_i = c_\alpha h^\alpha{}_b \quad , \tag{3}$$

where

$$c_\alpha = \mathbf{x}.\mathbf{q}_\alpha \tag{4}$$

is the scalar product of the vectors **x** and $\mathbf{q}_\alpha$. If $\mathbf{q}_\beta$ is the stored vector closest to the input **x**, then c_β is the largest among the coefficients c_α. Suppose that behind the BLT there is a stage which selects, from the linear combination $c_\alpha h^\alpha{}_b$, the dominant Hadamard vector $\mathbf{h}^\beta$. Such a device is here called a "*Dominant Label Selector* (DLS). We postpone discussion of the DLS, and consider the processing of the DLS output **y**, for now assumed to be the dominant Hadamard vector $\mathbf{h}^\beta$. As depicted schematically in Fig. 1, the DLS output **y** is returned to the BLT, to be used in a backstroke

$$w^i = y^b B_b{}^i \quad . \tag{5}$$

With $\mathbf{y} = \mathbf{h}_\beta$ and and the connection matrix (2), (5) gives as result of the BLT backstroke

$$w^i = h_\beta{}^b h_b{}^\alpha q_\alpha{}^i = N\, \delta_\beta{}^\alpha q_\alpha = N\, q_\beta{}^i \quad , \tag{6}$$

where use has been made of the orthonormality of the Hadamard vectors, expressed by (A2) in Appendix A. If **w** is thresholded with the signum function s_0 we get

$$x'^i = s_0(w^i) = q_\beta{}^i \quad , \tag{7}$$

which is the stored vector that is closest to the input **x**. Hence, if the DLS would work as

required, the SRM would perform perfect associative recall of anyone of N stored vectors. There would not be any spurious stable states.

There is the option of deleting the thresholding in front of the BLT. Then, no BLT neurons are needed; the BLT is just a "bidirectional connection box". Also, there is the option of using an analog input for **x**. Finally, there is the option of using the result of the BLT backstroke to upgrade the SRM input, or just as the output of the SRM.

The DLS has the task to select, from the linear combination $u_b = c_\alpha h^\alpha{}_b$, the Hadamard vector $h^\beta{}_b$ which occurs with the largest coefficient, i.e., c_β is largest among the c_α, α=1 to N. But this means that $\mathbf{h}_\beta$ is the Hadamard vector with the largest scalar product $\mathbf{u}.\mathbf{h}_\alpha$. Therefore, the DLS itself may be considered an associative memory with stored states $\mathbf{h}_\alpha$, α=1 to N. In those terms the DLS is to produce, from the input **u**, the closests stored state, $\mathbf{h}_\beta$.

DOMINANT LABEL SELECTOR (DLS)

The DLS, considered as an associative memory with stored states taken as the Hadamard vectors $\mathbf{h}_\alpha$, must find the Hadamard vector nearest to the BLT output **u**. A Hopfield memory [1] cannot be used here, because it would saturate long before all the N states are stored. Furthermore there is a problem due to orthonormality of the stored states [9]. Instead, we have chosen for the DLS an associative memory with quadratic activation. This memory, called "Quadratic Hadamard Memory", has been investigated in [9] by means of the asynchronous discrete model, in which the activation is given by

$$v_a = S_{abc} y^b y^c + r_a \quad , \tag{8}$$

and the signal y_a is determined by thresholding v_a with the signum function. The last term r_a may be either seen as an external coupling, or as defining thresholds. The connection tensor S_{abc} is restricted to be fully symmetric.

The quadratic activation expressed by (8) constitutes a case of "higher–order neurons" [11]. Properties of quadratic activations have recently been discussed by Volper and Hampson [12].

Stability of the DLS in the asynchronous discrete model is assured [11] if all connection tensor components with at least two equal indices are zero,

$$S_{app} = 0 \text{ , for all a and p.} \tag{9}$$

It has been shown in [9] that the Hadamard states $\mathbf{h}_\alpha$, α=1 to N, are stable in the asynchronous discrete model if the connection tensor is chosen as

$$S_{abc} = \sum_\alpha h_{\alpha a} h_{\alpha b} h_{\alpha c} - N\,\delta_{ab}\delta_{c1} - N\,\delta_{bc}\delta_{a1} - N\,\delta_{ca}\delta_{b1} + 2N\,\delta_{a1}\delta_{b1}\delta_{c1}. \tag{10}$$

and if, moreover, we take

$$r_a = r \quad \text{for all indices a.} \tag{11}$$

and $$0 < r < N^2-2N \, . \tag{12}$$

The form of the connection tensor (10) is similar to the Hopfield matrix [1], making allowance for the quadratic nature of the internal coupling. The last four terms in (10) have been added to satisfy condition (9), while retaining full symmetry. To check that the subtraction works one needs the property (A4) of the Hadamard vectors used.

It further has been shown in [9] that in the asynchronous discrete model no spurious stable states exist if one takes (10) and (11), and if (12) is sharpened to

$$N^2-6N < r < N^2-2N \, . \tag{13}$$

In addition to the system with the subtracted connection matrix (10), we considered in [9] an alternate system, in which the connection matrix is simply taken as

$$S_{abc} = \sum_{\alpha} h_{\alpha a} h_{\alpha b} h_{\alpha c} \, , \tag{14}$$

without any subtractions. In the asynchronous discrete model this DLS has been found [9] to also have the properties mentioned above. Although this system has a simpler expression for the connection matrix, it has a somewhat larger number of physical connections, because (9) does not hold.

CONTINUOUS MODEL FOR DLS

The main problem remaining after the discrete model investigation [9] was the question of coupling of the BLT output u_a to the DLS, taken as a quadratic Hadamard memory. The discrete model cannot properly account for such coupling, and this is the main reason for extending the investigation to the continuous model. The DLS dynamics is then described by equations of motion for the activation v_a of neuron a,

$$\dot{v}_a = -v_a + S_{abc}\, y^b y^c + r_a \, . \tag{15}$$

The dot denotes a time derivative, and the output signal y_a of neuron a is given in terms of the activation v_a by

$$y_a = s(v_a) \, , \tag{16}$$

where s(v) is a soft sigmoid function which either attains the values ± 1, or approaches these values asymptotically. The function is chosen antisymmetric and such that

$$s'(v) \geq 0 \quad \text{for all } v \, , \tag{17}$$

where the prime denotes the derivative.

We need to discuss the coefficients of terms in (15). In electronic implementations, the first two terms represent the lumped effects of amplifier input capacitance and resistance. This may be expressed more clearly by writing the terms as $C\dot{v}_a$ and $-v_a/R$. But C and R can be brought to unity by scaling of the time and the activation, together with a related adjustment of the sigmoid function s. Hence, putting C and R to unity does

not constitute a physical restriction. We will proceed with the DLS dynamics in the normalized form (15).

The connection tensor S_{abc} in the equations of motion needs to be specified. Two choices will be considered: the subtracted tensor (10) and the unsubtracted tensor (14). In the asynchronous discrete model these two connection tensors give about the same results [9]. In the sequel we will sometimes choose the subtracted form, and sometimes the subtracted tensor, as determined by opportunities for theory development or clarification. There also are results which hold for any symmetric connection tensor, and they will of course be derived without having the tensor specified.

Because in the continuous model the set of signal states is the closed or open solid hypercube $J_N = [-1,1]^N$ or $J_N^{'} = (-1,1)^N$, there are many more possibilities for stationary states than in the discrete model, where the signals are constraint to lie at the corner points of the solid hypercube. Hence, the investigation of stationary and stable points of the continuous DLS involves a lot more territory. In addition, there is the question of where the DLS state will eventually settle, if started out at a suitable initial state. The nonlinearities in the DLS dynamics make it difficult to integrate the equations of motion. Our challenge is to get the required information without having to do the integration.

NEURON OUTPUT FUNCTIONS

Two convenient choices have been made for the neuron output function s(v). The first is the hyperbolic tangent,

$$s(v) = \tanh(gv) \,, \tag{18}$$

where g is the gain at zero. Since the asymptotic values ±1 are not attained, the set of signals is here the <u>open</u> solid hypercube $J_N^{'} = (-1,1)^N$.

The other output function used in this report is the *piecewise linear function*

$$\begin{aligned} s(v) &= -1 \quad \text{if} \quad v < -1/g \\ &= gv \quad \text{if} \quad |v| \le 1/g \,, \\ &= 1 \quad \text{if} \quad v > 1/g \,. \end{aligned} \tag{19}$$

This function attains the values ±1, so that the set of signals is the closed solid hypercube $J_N = [-1,1]^N$. In analytical work one has to watch the discontinuity in the derivative at $v = \pm 1/g$. Moreover, if any of the acitivations v_a exceeds $1/g$ in magnitude, then the state cannot be described unambiguously by the signal **y**, and one must use the activation **v**. Either sigmoid function has its advantages and disadvantages, and we will use one or the other, as convenient. Although one must be careful not to claim more than is proved, we expect the results derived to remain valid for other similar choices of output function. In the numerical computations performed, no difference was noticed when one function was used or the other, as long as the gains g were taken the same. Properties and consequences of the two output functions are shown in Appendix C.

THE ENERGY

For a fully symmetric connection tensor S_{abc}, the activation *velocity* vector $\dot{\mathbf{v}}$ is curl free in the space of signal vectors $\mathbf{y}$ (but not in the space of activations; see Appendix B). Hence, the integral

$$E = -\int_0^{\mathbf{y}} \dot{v}_a \, dy^a \tag{20}$$

does not depend on which path is taken from the origin to the point $\mathbf{y}$. E is a Liapunov function since

$$\dot{E} = -\sum_a (\dot{v}_a)^2 \, s'(v_a) \tag{21}$$

is nonpositive, by (17). From (20) one finds

$$E(\mathbf{y}) = -\tfrac{1}{3} S_{abc} y^a y^b y^c - r_a y^a + \Omega \, , \tag{22}$$

where

$$\Omega = \sum_a \varphi(v_a) \tag{23}$$

and

$$\varphi(v) = \int_0^{v} \xi \, s'(\xi) \, d\xi . \tag{24}$$

Although φ is defined in (24) as a function of v, the term Ω in (22) is taken as a function of y_a, by application of the inverse of the mapping $y=s(v)$. The inverse mapping is unique if condition (17) is replaced by

$$s'(v) > 0 \, , \quad \text{for all } v \, . \tag{25}$$

Since the *energy* E of (22) is bounded, and $\dot{E}$ is nonpositive, we have

Theorem 1: The continuous DLS subject to the conditions posed is stable.

SPECIFIC FORMS FOR EQUATIONS OF MOTION AND ENERGY

For the subtracted connection tensor(10) we have from (15) the equations of motion

$$\dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha{}^2 - 2N y_a y_1 - N y^b y_b \delta_{a1} + 2N y_1{}^2 \delta_{a1} + r_a \, . \tag{26}$$

where

$$y_\alpha = (1/\sqrt{N}) h_{\alpha a} y^a \tag{27}$$

is the Hadamard transform of y_a. The factor $1/\sqrt{N}$ is applied in order to preserve norms. In (26), y_1 is the component y_a for $a=1$, not the component y_α for $\alpha=1$.

It is sometimes convenient to see separately the components for $a=1$ and for $a\neq 1$:

$$\dot{v}_1 = -v_1 + r_1 \,, \tag{28}$$

and

$$a \neq 1, \qquad \dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha{}^2 - 2N y_a y_1 + r_a \,; \tag{29}$$

use has been made of (A6) and (A9). The energy is

$$E = -\tfrac{1}{3}\left(\sqrt{N^3} \sum_\alpha y_\alpha{}^3 - 3N y^a y_a y_1 + 2N y_1{}^3\right) - r^a y_a + \Omega \,, \tag{30}$$

where Ω is given by (23) and (24). If the neuron output function is a hyperbolic tangent as given by (18), then one has for Ω

$$\Omega = \frac{1}{2g} \sum_a \{(1+y_a)\ln(1+y_a) + (1-y_a)\ln(1-y_a)\} \,, \tag{31}$$

as derived in Appendix C. If the output function is taken as the piecewise linear function (19), then

$$\Omega = R^2/(2g) \,, \text{ for } |y_a| < 1 \,, \tag{32}$$

where R is the Euclidean norm of $\mathbf{y}$,

$$R^2 = y^a y_a \,. \tag{33}$$

For the unsubtracted connection tensor (14) the equations of motion are

$$\dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha{}^2 + r_a \,. \tag{34}$$

which split into

$$\dot{v}_1 = -v_1 + N\, y^a y_a + r_1 \,, \tag{35}$$

and

$$a \neq 1, \qquad \dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha{}^2 + r_a \,. \tag{36}$$

where, again, (A6) and (A9) have been used. The energy is then

$$E = -\tfrac{1}{3}\sqrt{N^3} \sum_\alpha y_\alpha{}^3 - r^a y_a + \Omega \,. \tag{37}$$

For subtracted dynamics, the a=1 equation of motion (28) is uncoupled from the rest. The solution is simply

$$v_1 = r_1 + c e^{-t} \,, \tag{38}$$

where c is a constant. Since all Hadamard vectors used here have + as first component, and we want a Hadamard vector as DLS output, things must be arranged such that y_1 approaches unity for large times, which means that the activation v_1 must also be positive. It follows that we must require

$$r_1 > 0 \,. \tag{39}$$

For unsubtracted dynamics, the a=1 equation of motion is given by (35). Now there is coupling to the scalar R^2 of (33). The condition (39) then assures that v_1 does not

temporarily turn negative, which would slow the settling of the DLS.

For subtracted dynamics there is the option of omitting neuron #1 altogether, and clamping the y_1 signal line to $y_1 = +$ permanently. There is no need for connections from the y_1 line to the other neurons. We will refer to this arrangement as the *y_1 clamping arrangement.* This amounts to putting

$$y_1 = + \,, \tag{40}$$

which gives the $-2Ny_a y_1$ term in (29) the value $-2Ny_a$. The first component, u_1, of the BLT output u_a is ignored. That this can be done without penalty is related to the fact that the Hadamard vectors used all have first component +. The y_1 clamping speeds up the DLS action, as can be seen by monitoring numerical computations.

For unsubtracted dynamics, the coupling term $N\, y^a y_a$ in (35) is always positive, and it tends to N^2 or a nearby value for large times. The value of r_1 needs to be chosen such that y_1 tends to +1 for large times. The y_1 clamping scheme may be used also for the DLS with unsubtracted dynamics.

COUPLING SCHEMES

The BLT output **u** must be coupled to the DLS. We see two ways of doing this. In the *external* coupling scheme the vector **r** in the equations of motion is chosen as

$$r_a = \mu(u_a + cN\delta_{a1}) \,, \tag{41}$$

where $\mu>0$ is a coupling constant, and N is the dimension. For a=1 the Kronecker δ_{a1} is unity, else zero. The term $cN\delta_{a1}$ applies a threshold $-\mu cN$ solely to the first neuron. This threshold has been written into **r** for later convenience, and the constant c will be chosen in due time. The external coupling scheme also requires a reset of the DLS activation to zero, everytime a new BLT output **u** is applied. In the theory, such time is chosen as t=0, and the reset then fixes an initial value for the activation vector **v**(t):

$$v_a(0)=0 \,, \forall\, a \,. \tag{42}$$

For small times the quadratic term in the equations of motion (15) is negligible compared to r_a , so that we have

$$t<<1 \,, \quad \dot{v}_a = -v_a + r_a \,, \tag{43}$$

with r_a given by (41). With the initial value (42) we have the integral

$$t<<1, \;\; v_a = \mu(u_a + cN\delta_{a1})(1-e^{-t}) \,, \tag{43}$$

which shows that the activation exponentially approaches $\mu(u_a + cN\delta_{a1})$. The e folding time is unity here, because the RC time is unity, by the scaling that has thrown the

equations of motion into the normalized form (15).

The second choice is the *initial value* coupling, for which the vector **r** is chosen as

$$r_a = c\delta_{a1} , \tag{44}$$

with c some constant to be determined, and the BLT output is applied as initial activation, after multiplication by the coupling constant μ,

$$v_a(0) = \mu u_a . \tag{45}$$

For small times we now have the integral

$$t << 1 , \; v_a = cN\delta_{a1}(1 - e^{-t}) + \mu u_a e^{-t} . \tag{46}$$

The activation components now approach the value $cN\delta_{a1}$, and the contribution of the applied BLT output **u** in the activation dies out.

Comparing the two coupling schemes, the external coupling appears to have the practical advantage that the DLS input and output are separate. The BLT output vector **u** remains standing on the DLS external input, while on the output the DLS state **y** appears as it is developing in time. The output **y** is processed by the BLT in its backstroke, with the result **x′** appearing at the front of the BLT. The separation of the DLS input and output is particularly convenient for the setup in which the BLT front output **x′** is not used to upgrade the input **x** , but is considered as the output of the whole machine.

DYNAMIC REGIONS

It is helpful to distinguish regions in signal space which have essentially different dynamics. These regions do not have sharp boundaries, but blend smoothly into each other. In discussing these regions, we prefer to use simple albeit imprecise language rather than cumbersome precision.

In the *proportional region* of signal space the $s'(v_a)$ is constant, either precisely or approximately. This region includes the origin, and it may have considerable extent, depending on the neuron output function s(v) used. For the piecewise linear function (19) the proportional region in signal space is given by $-1 < y_a < 1$.

In the proportional region $y_a = gv_a$, where g is the constant gain in the region, and the equations of motion (15) may be written

$$\dot{y}_a = -y_a + gS_{abc}y^b y^c + gr_a . \tag{47}$$

For unsubtracted dynamics the equations of motion (47) can be decoupled by means of a Hadamard transform; with (14) one finds

$$\frac{1}{g}\dot{y}_\alpha = -\frac{1}{g}y_\alpha + \sqrt{N^3}\, y_\alpha^{\,2} + r_\alpha , \tag{48}$$

where y_α is the Hadamard transform (27) of y_a and r_α is the Hadamard transform of r_a.

For subtracted dynamics one finds with (10)

$$\frac{1}{g}\dot{y}_\alpha = -\frac{1}{g}y_\alpha + \sqrt{N^3}\, y_\alpha{}^2 - 2N\, y_\alpha y_1 - \sqrt{N}\, y^\beta y_\beta h_{\alpha 1} + 2\sqrt{N}\, y_1{}^2 h_{\alpha 1} + r_\alpha , \qquad (49)$$

where (A6) has been used. In (49) and throughout this report, y_1 stands for y_a with a=1. If y_1 clamping is used, y_1=1 is to be substituted in (49). Even then, the equations of motion are not entirely uncoupled, because of the term with

$$y^\beta y_\beta = R^2 . \qquad (50)$$

Note that the norms given by (50) and (33) are equal, because of the orthogonality of the Hadamard Transform.

The neuron output function s(v) is restricted to be of the sigmoid type. For estimation purposes it is convenient to introduce a value v^*, which we call the *critical activation*, such that

$$\text{for } |v| > v^* , \quad |s(v)| \geq 1 - \epsilon = y^* , \qquad (51)$$

where ϵ is a positive number much smaller than unity, such as $\epsilon = 0.01$. Condition (17) or (25) and the antisymmetry of s(v) imply

$$\text{for } |v| > v^* , \; s'(v) \leq s'(v^*) . \qquad (52)$$

We call an activation v_a *subcritical* if $|v_a| < v^*$, and *supercritical* if $|v_a| \geq v^*$. For a supercritical activation v_a, the resulting signal y_a may be taken as ± 1, determined by the sign of v_a, as a suitable approximation in certain mathematical expressions. If (25) is true, sub– and supercriticallity also can be stated in signal space: a signal y_a is subcritical if $|y_a| < y^*$, and supercritical if $|y_a| \geq y^*$. Hence, in discussing sub– and supercriticallity, we then need not say whether the state is considered in activation space or signal space.

Since the state vector **v** has N components, some components may be supercritical , while others may be subcritical. Hence we distinguish states that are *entirely subcritical*, *partially supercritical*, and *entirely supercritical*. A supercritical activation not only produces a signal that may be approximated as ± 1, but also gives a derivative s′ smaller than $s'(v^*)$, by (52). Which if any of these two properties of supercriticallity is used is a matter of convenience.

When the state is entirely supercritical, the point **y** lies at or close to a corner point of the solid hypercube. The point set $\{\mathbf{y} | y^* < |y_a| \leq 1\}$, with a suitable y^* is called a *corner*. If a Hadamard point is included, we call the set a *Hadamard corner*.

STATIONARY POINTS OF THE I_N

It is a simple matter to determine whether a given point **y** is stationary. All one

needs to do is calculate $\dot{\mathbf{v}}$ from the equations of motion (26) or (34), and see whether the activation velocity $\dot{\mathbf{v}}$ is zero. Alternatively, one can see whether at $\mathbf{y}$ the energy E has a zero gradient,

$$\partial_a E = 0 , \tag{53}$$

where ∂_a stands for $\partial/\partial y^a$.

As will be shown, thresholds can be chosen such that the Hadamard points are stationary. The question is whether there are other stationary points. To investigate this in the continuous model we use a method suggested by the study [9] of the asynchronous discrete model. We begin by considering the vector

$$Q_a = N \sum_\alpha h_{\alpha a} y_\alpha^2 = \sum_\alpha h_{\alpha a} h_{\alpha b} h_{\alpha c} y^b y^c , \tag{54}$$

in the equation of motion (29) for the case with the subtracted connection tensor (10);

$$a \neq 1, \qquad \dot{v}_a = -v_a + Q_a - 2N y_a y_1 + r_a . \tag{55}$$

We choose the piecewise linear output function given by (19), the gain must be chosen large enough so that some stationary points belong to I_N. The vector Q_a can be rewritten with two Hadamard factors by using the group property of Hadamard vectors, (see Appendix A)

$$h_{\alpha a} h_{\alpha b} = h_{\alpha d} , \tag{56}$$

where

$$d = f(a,b) ; \tag{57}$$

this allows rewriting (54) as

$$Q_a = \sum_\alpha \sum_{b,c} h_{\alpha d} h_{\alpha c} y_b y_c . \tag{58}$$

Using the group property once more,

$$h_{\alpha d} h_{\alpha c} = h_{\alpha e} , \tag{59}$$

where

$$e = f(d,c) ; \tag{60}$$

throws (58) in the form

$$Q_a = \sum_\alpha \sum_{b,c} h_{\alpha e} y^b y^c . \tag{61}$$

With (A4), (61) may be written

$$Q_a = N \sum_{b,c} \delta_{e1} y^b y^c . \tag{62}$$

In order to contribute to the sum, the term $\delta_{e1} y^b y^c$ must have e=1; with (60) and (A17) this means that d=c . With (57) and (A22) this implies that

$$a = f(b,c) . \tag{63}$$

Hence, (62) may be written

$$Q_a = N \sum_{(b,c)\in\aleph_a} y^b y^c , \tag{64}$$

where $\aleph_a$ is the set of index pairs

$$\aleph_a = \{(b,c) |\ f(b,c)=a\} . \tag{65}$$

We restrict to signals that lie on cornerpoints of the solid hypercube, i.e., $\mathbf{y}\epsilon I_N$.

As further preparation, we state here as Lemma 1 the Theorem 2 of [9]; for convenience, the proof [9] is repeated in Appendix D.

Lemma 1: $\mathbf{y}\epsilon I_N$, $Q_b=-N^2$, $\forall$ b such that $y_b=-$ <==> $\mathbf{y}$ is Hadamard $\neq \mathbf{h}_1$.

We are now prepared to show

Lemma 2: $\mathbf{y} \epsilon I_N$, not Hadamard ==> $\exists$ index b such that $y_b = -$ and $Q_b \geq -N^2+4N$.

Proof: Let $\mathbf{y} \in I_N$ and not Hadamard. This implies that $\mathbf{y}$ is not $\mathbf{h}_1$, so that the set $A = \{b | y_b = -\}$ is not empty. Choose an index b such that $y_b=-$ and $Q_b \neq -N^2$; this is always possible, since otherwise Lemma 1 would imply that $\mathbf{y}$ is Hadamard $\neq \mathbf{h}_1$, which is false. We use the expression (64) for Q_a. There are N terms in the sum, since for every b=1 to N the remaining index c is determined by the condition that the pair (b,c) lies in the set $\aleph_a$. Since $\mathbf{y} \epsilon I_N$, all terms $y^b y^c$ are either 1 or –1. The terms cannot all be –, since $Q_a \neq -N^2$. Hence, the sum contains at least one + term. Say, that term is $y^p y^q$. But then, the term $y^q y^p$ is + as well. It follows that the sum (64) contains at least two + terms. Since flipping the sign of a single term from – to + causes the sum to change by 2, we have

$$Q_a \geq -N^2+4N. \qquad \rfloor$$

Lemma 2 has an important application to the equations of motion (55), with the threshold term chosen as

$$a\neq 1, \qquad r_a = (N^2-4N)h_{1a}+\mu u_a , \tag{66}$$

where $\mathbf{h}_1$ is the Hadamard vector with all components +, u_a is the output of the BLT, and $\mu > 0$ is a coupling constant. With (66), the BLT is coupled to the DLS by applying the BLT output as an external coupling to the DLS. It is convenient to put a bound on the coupling constant,

$$a\neq 1, \quad |\mu u_a| < 2N-1/g . \tag{67}$$

and restrict the gain g by

$$g > \frac{1}{2N} . \tag{68}$$

Moreover we take from here on

$$N>4. \tag{69}$$

We have

Theorem 2: For a continuous DLS with subtracted connection tensor (10), a piecewise linear output function, thresholds subject to (66) and (67), and a gain subject to (68), the only stationary points of I_N are the Hadamard points.

Proof: First we show that the Hadamard points are stationary. For a Hadamard point $\mathbf{y} = \mathbf{h}_\gamma$ one has

$$Q_a = N^2 h_{\gamma a}\,, \tag{70}$$

and the equation of motion (55) with thresholds given by (66) reads

$$a\neq 1, \qquad \dot{v}_a = -v_a + N^2 h_{\gamma a} - 2N h_{\gamma a} + N^2 - 4N + \mu u_a\,. \tag{71}$$

where $y_1 = +$ has been substituted.

For indices a such that $h_{\gamma a} = -$ the N^2 terms cancel, and (71) gives

$$\dot{v}_a + v_a = -2N + \mu u_a < -1/g\,, \tag{72}$$

while for the remaining indices $a\neq 1$ we have

$$\dot{v}_a + v_a = 2N^2 - 6N + \mu u_a > 1/g\,; \tag{73}$$

conditions (67), (68), and (69) have been used. (72) shows $\exists\, v_a < -1/g$ such that $\dot{v}_a = 0$; the inequality $v_a < -1/g$ is consistent with $y_a = -1$. (73) shows $\exists\, v_a > 1/g$ such that $\dot{v}_a = 0$; $v_a > 1/g$ is consistent with $y_a = 1$. The a=1 equation of motion (28),

$$\dot{v}_1 = -v_1 + r_1 \tag{74}$$

shows no coupling with other neurons, and can either be implemented as is, with $r_1 > 0$, or may be cast aside in favor of y_1 clamping, as discussed before.

It follows that the Hadamard point $\mathbf{y} = \mathbf{h}_\gamma$ is stationary. To show that there are no other stationary points, consider a point $\mathbf{y}$ of the I_N that is not Hadamard. By Lemma 2 there exists an index b such that $y_b = -$ and $Q_b \geq -N^2 + 4N$. For such an index, the equation of motion (55) gives

$$\dot{v}_b + v_b = Q_b + 2N + N^2 - 4N + \mu u_b \geq 2N + \mu u_b > 0. \tag{75}$$

The point cannot be stationary because of a conflict in signs of v_b and y_b. ⌋

Condition (69) does not constitute a restriction in practical applications. The gain conditions (68) is satisfied in practice because we want large gains in order that the DLS settles fast. With (69), condition (67) is satisfied if

$$\mu<1\,, \tag{76}$$

because

$$|u_a| \leq N\,, \tag{77}$$

as follows from (3), $c_\alpha = \mathbf{x}\cdot\mathbf{q}_\alpha$, and the bipolar nature of the vectors $\mathbf{x}$ and $\mathbf{q}_\alpha$. Condition (76) is convenient in practice, since it implies that no amplification is needed between the BLT rear output and the DLS input.

STABILITY

The stability of stationary points can be investigated either with the energy or with the equations of motion.

Writing ∂_a for $\partial/\partial y^a$, we have for unsubtracted dynamics, from (37)

$$\partial_a E = v_a - N \sum_\alpha h_{\alpha a} y_\alpha^{\,2} - r_a\,, \tag{78}$$

and

$$\partial_b \partial_a E = \delta_{ba}/s'(v_b) - 2\sqrt{N} \sum_\alpha h_{\alpha a} h_{\alpha b} y_\alpha\,. \tag{79}$$

Since away from stable points $\dot{E}$ of (21) is negative definite, the stationary point $\mathbf{y}$ is asymptotically stable iff the tensor $\partial_b \partial_a E$ at $\mathbf{y}$ is positive definite. The point is unstable iff the tensor has a negative eigenvalue. It is difficult to use these conditions because of the involvement of the Hadamard matrices in the tensor (79). But, if we are willing to give up some bound sharpness, a very simple condition can be stated in terms of the tensor trace; if it is negative, then there must be a negative eigenvalue. Hence $\partial^a \partial_a E < 0$ implies that the stationary point is unstable. With (79), (A2), and (A4) it follows after a short calculation that

$$\sum_a (1/s'(v_a)) - 2N^2 y_1 < 0 \Longrightarrow \mathbf{y} \text{ is unstable.} \tag{80}$$

This sufficient condition for instability only involves y_1 and the sum of the reciprocal sigmoid derivatives s' for the neurons. Since s' is nonnegative by (17) or (25), cancellations cannot occur in the sum over a. Therefore, satisfaction of the condition requires that <u>none</u> of the derivatives $s'(v_a)$ be small. Roughly, this means that the state must be entirely subcritical in order that instability can be concluded from (80).

With about the same effort a much sharper sufficient condition for instability can be derived from the equations of motion. These may be seen as expressing the flow velocity $\dot{\mathbf{v}}$ in activation space. The flow divergence is related to instability, as will be shown presently.

Writing ∂_a for $\partial/\partial v^a$, we have from (34)

$$\partial_b \dot{v}_a = -\delta_{ba} + 2\sqrt{N} \sum_\alpha h_{\alpha a}\, y_\alpha\, h_{\alpha b}\, s'(v_b) \tag{81}$$

At a stationary point $\mathbf{v}$ the velocity $\dot{v}_a$ vanishes. The velocity at a point $\delta\mathbf{v}$ away from $\mathbf{v}$ is

given by $\delta v^b \partial_b \dot{v}_a$. The question is whether the radial component of this velocity is pointing away from the point $\mathbf{v}$ or towards it. If the former is the case for some vector δv^b, the point $\mathbf{v}$ is unstable; if the latter occurs for every vector δv^b, then the point is stable. The sign of the radial component of the vector $\delta v^b \partial_b \dot{v}_a$ is given by the sign of the scalar product $\delta v^a \delta v^b \partial_b \dot{v}_a$. It follows that $\mathbf{v}$ is unstable if the tensor $\partial_a \dot{v}_b + \partial_b \dot{v}_a$ has a positive eigenvalue. This is the case if the trace is positive. That trace is the flow *divergence*

$$\partial_a \dot{v}^a = -N + 2Ny_1 \sum_a s'(v_a) . \tag{82}$$

We have

$$\partial_a \dot{v}^a > 0 \implies \mathbf{v} \text{ is unstable,} \tag{83}$$

and with (82) the condition reads

$$1 - 2y_1 \sum_a s'(v_a) < 0 \implies \mathbf{v} \text{ is unstable.} \tag{84}$$

As in (80), the condition involves only y_1 and the derivatives s'. However, (84) is much sharper than (80): all that is required for the satisfaction of the inequality is that for a single neuron the sigmoid derivative s' is sizeable. Roughly, this means that partially or entirely subcritical stationary states are unstable. A further advantage of the condition (84) over (80) is the absence of the factor N^2, which is very large for the large dimensions expected to be important in practice.

We proceed with application of (84) to the case that the neuron output function is taken as the hyperbolic tangent (18). Then we have from the Appendix, (C3),

$$s'(v_a) = g(1-s^2(v_a)) = g(1-y_a^2) . \tag{85}$$

The condition (84) then reads

$$1-2gy_1 \sum_a (1-y_a^2) < 0 \implies \mathbf{y} \text{ is unstable} . \tag{86}$$

For partially or entirely subcritical signals $\mathbf{y}$

$$\exists\, b \text{ such that } |y_b| < y^* , \tag{87}$$

where

$$y^*=s(v^*) , \tag{88}$$

and v^* is a suitably chosen critical activation. The inequality in (86) is satisfied if

$$0<\eta<1, \quad 2g\eta(1-y^{*2}) > 1, \; y_1 > \eta , \text{ and (87) is true;}$$

Since for $0<\epsilon<1$ and $y^*=1-\epsilon$ we have $1-y^{*2} > \epsilon$, it follows that for partially or entirely subcritical signals

$$2g\eta\epsilon>1, \; y_1>\eta, \; |y_a|<y^* \implies \mathbf{y} \text{ is unstable} . \tag{89}$$

The number η may be chosen freely, as long as $0<\eta<1$. A convenient choice is

$$\eta = 1/2 \; ; \tag{90}$$

then, the condition $2g\eta\epsilon>1$ in (89) becomes

$$\epsilon>1/g \; . \tag{91}$$

Hence, we have

Theorem 3: For a DLS with unsubtracted dynamics and a hyperbolic tangent output function with gain g, any stable point must lie either in the region $y_1 \leq 1/2$, or in the entirely supercritical region with criticallity parameter $\epsilon>1/g$.

As will be shown later, the only stable point in the region $y_1 \leq 1/2$ is at or near the origin, and it can be eliminated by proper choice of coupling constant μ. By Theorem 3, the remaining stable points must lie in the corners of the solid hypercube with positive y_1. Of course, the Hadamard corners are of special interest; we want to know whether they contain a stable point. Theorem 2 states that Hadamard points are stationary if certain conditions are satisfied, which includes the threshold condition (66). But it turns out that the large magnitude of the threshold in (66) spoils the early dynamics, and therefore we will need to diminish the threshold below the value given by (66). In preparation, we must find a range of thresholds that straddles the origin, and which assures that every Hadamard corner contains a stable point. The first step is to find a threshold range such that every Hadamard corner contains a stationary point. We proceed as follows.

For subtracted dynamics we have the equations of motion (29)

$$a\neq 1, \qquad \dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha^{\,2} - 2N y_a y_1 + r_a \tag{92}$$

We wish to find bounds on the threshold such that the Hadamard points are stationary. If we use the hyperbolic tangent output function (18), the Hadamard points are not attainable; we then consider the signal

$$y_a = (1-\epsilon_a) h_{\gamma a}, \tag{93}$$

where

$$0<\epsilon_a<.02 \; , \; \forall \; a. \tag{94}$$

The value .02 has been chosen for convenience in a manner that need not be discussed here. For this signal the equation of motion (92) gives

$$a\neq 1, \; \dot{v}_a = -v_a + N^2(1-\epsilon_a)^2 h_{\gamma a} - 2N(1-\epsilon_a) h_{\gamma a}(1-\epsilon_1) + r_a \; . \tag{95}$$

We investigate whether the signal can be stationary. Putting $\dot{v}_a = 0$ gives

$$a\neq 1, \; v_a = N(1-\epsilon_a)(N(1-\epsilon_a) - 2(1-\epsilon_1)) h_{\gamma a} + r_a \; . \tag{96}$$

If

$$|r_a| < N^2/4 \; , \tag{97}$$

then the sign of v_a as determined by (96) is the same as the sign of y_a given by (93), since

$$N(1-\epsilon_a)(N(1-\epsilon_a) - 2(1-\epsilon_1)) - N^2/4 > 0 \tag{98}$$

for all ϵ_a subject to (94). It remains to calculate ϵ_a such that the signal **y** given by (93) is

entirely supercritical, with criticallity parameter ϵ_a. For r_a subject to (97), Eq. (96) implies

$$a\neq 1,\ |v_a|>N(1-\epsilon_a)(N(1-\epsilon_a)-2)-N^2/4\ , \tag{99}$$

provided that ϵ_a has been chosen such that

$$\epsilon_a<\epsilon\ ,\ \forall\ a\ . \tag{100}$$

Let g^* be determined such that

$$N(1-\epsilon)(N(1-\epsilon)-2)-N^2/4=\frac{1}{2g^*}\ln(2/\epsilon)\ . \tag{101}$$

(99) can then be written

$$a\neq 1\ ,\ |v_a|>\frac{1}{2g^*}\ln(2/\epsilon)>\frac{1}{2g}\ln(2/\epsilon)\ , \tag{102}$$

for all gains g such that

$$g>g^*\ . \tag{103}$$

By (C6) we have

$$\frac{1}{2g}\ln(2/\epsilon)>v^*\ , \tag{104}$$

where v^* is the critical activation belonging to ϵ, as defined by (C4). It follows that (102) implies

$$a\neq 1,\ |v_a|>v^*\ . \tag{105}$$

For a=1 we have (28) , and at the stationary point the activation v_1 can be made as large as desired by choosing r_1 large enough. Alternatively, one can use the y_1 clamping scheme and put $y_1=+$. Together with (105) and (100) it follows that the state (93) is entirely supercritical.

The bound (97) is sloppy, but very generous for practical applications. We must see whether the gains g subject to (103) and (101) have practical values. For N subject to (69) it was found that

$$.02<\epsilon\leq.001\ ,\ g^*<1\ . \tag{106}$$

Since we want large gains for fast DLS settling, the condition (103) does not constitute a limitation in practice.

It follows that a DLS with subtracted dynamics and hyperbolic tangent output function has a stationary point in every Hadamard corner, provided that (97) is satisfied.

For unsubtracted dynamics the argument is much the same. The term $-2Ny_ay_1$ is then missing from (92), so that instead of (96) we have

$$a\neq 1,\ v_a=N^2(1-\epsilon_a)^2h_{\gamma a}+r_a \tag{107}$$

In this case, we need only invoke (101); the same bound (97) then assures the existence of stationary point in every Hadamard corner, for a DLS with unsubtracted dynamics.

It is easy to see that the conclusions remain valid if the hyperbolic tangent output function is replaced by the piecewise linear function given by (19), for both dynamics

considered. Of course, since the outputs ±1 are attainable, the stationary point in a Hadamard corner lies precisely at the Hadamard point.

We have shown

Theorem 4: For a DLS with either subtracted or unsubtracted dynamics, with the output function taken either as a hyperbolic tangent or as a piecewise linear function, there exists a stationary point in every Hadamard corner, if the bound (97) on thresholds is satisfied, and the gain is at least the g^* determined from (101).

Next, we investigate the stability of these stationary points in Hadamard corners. From (79) we have

$$\delta^2 E = \sum_a \delta y_a{}^2 / s'(v_a) - 2\sqrt{N^3} \sum_\alpha \delta y_\alpha{}^2 y_\alpha , \tag{108}$$

where $\delta^2 E = \delta y^b \delta y^a \partial_b \partial_a E$ is the second variation of the energy due to a displacement $\delta \mathbf{y}$ (the first variation vanishes since $\mathbf{y}$ is stationary).

For the hyperbolic tangent output function (18) one has (see Appendix, (C3))

$$s'(v_a) = g(1-s^2(v_a)) = g(1-y_a{}^2) , \tag{109}$$

and (108) becomes

$$\delta^2 E = \sum_a \delta y_a{}^2 / (1-y_a{}^2) - 2\sqrt{N^3} \sum_\alpha \delta y_\alpha{}^2 y_\alpha . \tag{110}$$

Let the index b be such that $|v_b|$ is smallest among the $|v_a|$. Then $y_b{}^2$ is smallest among the $y_a{}^2$, and we have

$$1/(1-y_a{}^2) \geq 1/(1-y_b{}^2) , \ \forall\, a . \tag{111}$$

Hence, from (110) we have

$$\delta^2 E \geq \delta \mathbf{y}.\delta \mathbf{y}/(1-y_b{}^2) - 2\sqrt{N^3} \sum_\alpha \delta y_\alpha{}^2 y_\alpha . \tag{112}$$

The smallest rhs occurs when $\sum_\alpha \delta y_\alpha{}^2 y_\alpha$ is maximum, while $\delta \mathbf{y}$ is constraint to have fixed norm. In this regard we have

Lemma 3: Let there be an index β such that $y_\beta > y_\alpha$, $\forall\, \alpha \neq \beta$. For U_α subject to $U^\alpha U_\alpha = 1$, the maximum of $X = \sum_\alpha U_\alpha{}^2 y_\alpha$ then occurs at $U_\alpha = \delta_{\alpha\beta}$.

Proof: The stationary points of X, subject to

$$U^\alpha U_\alpha = 1 \tag{113}$$

are found from the stationary points of $F = X + \lambda(U^\alpha U_\alpha - 1)$, where λ is a Lagrangian multiplier. One has

$$0=\partial F/\partial U^{\alpha}=2U_{\alpha}y_{\alpha}-2\lambda U_{\alpha}\,. \tag{114}$$

λ is found by multiplication by U_{α}, summing over α, and using (113); the result is

$$\lambda=X\,. \tag{115}$$

Substitution in (114) gives

$$0=U_{\alpha}(y_{\alpha}-X)\,, \tag{116}$$

which implies

$$\text{either } U_{\alpha}=0\,, \text{ or } y_{\alpha}=X\,. \tag{117}$$

It follows that X is maximum for

$$U_{\alpha}=\pm\delta_{\alpha\beta}\,, \tag{118}$$

where β is the index such that $y_{\beta}>y_{\alpha}$, $\forall\ \alpha\neq\beta$. ⌋

For a stationary point in a Hadamard corner of the Hadamard vector $\mathbf{h}_{\beta}$ the maximum component y_{α} is y_{β}. Application of Lemma 3 then gives for the maximum of the term $2\sqrt{N^3}\sum_{\alpha}\delta y_{\alpha}{}^2 y_{\alpha}$ in (112), with $\delta\mathbf{y}$ subject to a fixed norm, the value

$$2\sqrt{N^3}\ \delta\mathbf{y}.\delta\mathbf{y}\ y_{\beta}<2N^2\ \delta\mathbf{y}.\delta\mathbf{y}\,. \tag{119}$$

Hence, (112) implies

$$\delta^2E\geq\delta\mathbf{y}.\delta\mathbf{y}\ (1/(1-y_b{}^2)-2N^2)\,. \tag{120}$$

The rhs is positive if

$$1/(1-y_b{}^2)>2N^2\,. \tag{121}$$

For ease of reference we state the result,

Lemma 4: For a DLS with unsubtracted dynamics the stationary points in Hadamard corners are stable if (121) is satisfied.

We must find a convenient inequality which implies (121). For the hyperbolic tangent output function $y=s(v)$ of (18) we have

$$1/(1-y^2)=\cosh^2(gv)=(e^{gv}+e^{-gv})^2/4\,; \tag{122}$$

hence, (121) is satisfied if

$$e^{g|v|}>2N\sqrt{2}\,. \tag{123}$$

where v is the activation with the smallest magnitude. (123) is equivalent to

$$|v|>\frac{1}{g}\ln(2N\sqrt{2})\,. \tag{124}$$

The equilibrium equations of motion (96) may be used to find a condition on r_a that

implies (124). From (96) we have for the minimum magnitude v of $|v_a|$

$$v > N(1-\epsilon)(N(1-\epsilon)-2) - |r_a| \; ; \tag{125}$$

with the bounds (94) and (97) the inequality (125) gives

$$v > .71N^2 - .98N \; . \tag{126}$$

Hence, (124) is satisfied if

$$.71N^2 - .98N > \frac{1}{g} \ln(2N/\sqrt{2}) \; . \tag{127}$$

With (69), the inequality is true for $g>1/4$. This can be seen by a direct calculation for $N=4$ and $g=1/4$, and the fact that the lhs of (127) increases more rapidly with N than the rhs. Hence, we have

Theorem 5: For a DLS with unsubtracted dynamics and hyperbolic tangent output function, the stationary point of Theorem 4 is stable if $N>4$, $g>1/4$, and $|r_a| < N^2/4$.

The conditions of Theorems 4 and 5 are easily satisfied in practice. Hence we have the result that our DLS has a stable point in each Hadamard corner, for a large range of thresholds that straddle the origin. Theorems 4 and 5 are in agreement with Theorem 3 for gains $g>50$. The bounds used for the derivation of Theorem 3 are very sloppy, and have been chosen in order to keep down the analytical work. (91) may be replaced by a less stringent condition by using tighter bounds. In any case, Theorems 4 and 5 by themselves suffice to assure a stable point in every Hadamard corner.

ENERGY LANDSCAPE

The energy provides a simple and natural way to visualize the dynamics. The energy is a scalar function in signal space, and by ignoring N–2 dimensions in some vague way we can imagine the energy function as a surface in three dimensional space. A point on the surface depicts the signal $\mathbf{y}$ as the projection on the horizontal plane, and the energy $E(\mathbf{y})$ as the height of the point. The dynamics drives the signal point down the energy surface, but generally not along the steepest path. This may be seen as follows. From (20) one has

$$\partial_a E = -\dot{v}_a = -\dot{y}_a / s'(v_a); \tag{128}$$

hence, in signal space the state changes in time by

$$\dot{y}_a = -(\partial_a E)/s'(v_a) \; . \tag{129}$$

Unless the derivatives $s'(v_a)$ have the same value for all a, the direction of the signal velocity $\dot{\mathbf{y}}$ differs from the direction of –grad E. In the proportional region, these directions are the same because $s'(v_a)=g$, $\forall$ a. Hence, in the proportional region, the signal point moves down the energy surface along the steepest path.

We explore the energy landscape for the case of maximum symmetry,

$$\mathbf{r} = 0 \; ; \tag{130}$$

application of a nonzero vector $\mathbf{r}$ would amount to tilting the energy surface, and this would cause a change in the stationary points that is easily visualized. With (130), the energy (22) is

$$E(\mathbf{y}) = -\frac{1}{3} S_{abc} y^a y^b y^c + \Omega \; . \tag{131}$$

We restrict signals to lie in the proportional region. With R the Euclidean norm of the signal, as given by (33), one then has

$$\Omega = R^2/(2g) \tag{132}$$

exactly for the piecewise linear output function, and approximately for the hyperbolic tangent, as shown in Appendix C. Hence, for either output function we have

$$E = -\frac{1}{3} S_{abc} y^a y^b y^c + R^2/(2g) \; . \tag{133}$$

A few results may be derived without specifying the form of the connection tensor S_{abc}, beyond symmetry. Writing ∂_a for $\partial/\partial y^a$, we have, from (133),

$$\partial_a E = - S_{abc} y^b y^c + y_a/g \; , \tag{134}$$

and

$$\partial_b \partial_a E = -2 \, S_{abc} y^c + \delta_{ba}/g \; . \tag{135}$$

At stationary states one has $\partial_a E=0$, so that (134) gives

$$y_a = g S_{abc} y^b y^c \; . \tag{136}$$

This equation has the solution

$$\mathbf{y}=0 \; , \tag{137}$$

and perhaps other solutions as well, which we denote by $\underset{r}{\mathbf{y}}$. The stationary point given by (137) is the origin; at that point, (135) gives

$$\partial_b \partial_a E = \delta_{ba}/g \; . \tag{138}$$

Since the tensor is positive definite, we have

Theorem 6: For a continuous DLS with $\mathbf{r} = 0$, and a neuron output function which is either a hyperbolic tangent or a piecewise linear function, the origin is asymptotically stable.

In order to investigate the stability of the other stationary points, $\underset{r}{\mathbf{y}}$, we resort to a trick that allows efficient use of equation (136). For a radial displacement

$$\delta y_a = \delta\mu\, y_a \tag{139}$$

one gets

$$\delta^2 E = \delta y^a \delta y^b \partial_b \partial_a E = (\delta\mu)^2(-2y^a S_{abc} y^b y^c + R^2/g)\,, \tag{140}$$

where (135) and (33) have been used. In the S term equation (136) can be used readily, with the result

$$\delta^2 E = (\delta\mu)^2(-2y^a y_a + R^2)/g\,. \tag{141}$$

Using (33) once more, (141) may be written as

$$\delta^2 E = -(\delta\mu)^2 R^2/g\,. \tag{142}$$

Since the second variation $\delta^2 E$ is negative, we have

Theorem 7: For a DLS with $\mathbf{r}=0$, and an output function that is either a hyperbolic tangent or a piecewise linear function, the stationary states $\underset{r}{\mathbf{y}}$ in the proportional region, but away from the origin, are unstable.

Calculation of the stationary points $\underset{r}{\mathbf{y}}$ requires that the connection tensor be specified. For the unsubtracted tensor (14), equation (136) takes the form

$$y_a = gN \sum_\alpha h_{\alpha a}\, y_\alpha^{\,2}\,. \tag{143}$$

The equations can be decoupled by means of a Hadamard transform, with the result

$$y_\alpha = g\sqrt{N^3}\, y_\alpha^{\,2}\,. \tag{144}$$

This implies that

$$\text{either } y_\alpha = \frac{1}{g\sqrt{N}3} \text{ or } y_\alpha = 0\,. \tag{145}$$

Denote by A the set of all indices α for which $y_\alpha \neq 0$,

$$A = \{\alpha |\ y_\alpha \neq 0\}\,. \tag{146}$$

By (145), the solutions $\underset{r}{\mathbf{y}}$ of (143) away from the origin may then be written

$$\text{if } \alpha \in A,\ \underset{r}{y}_\alpha = 1/(g\sqrt{N^3})\,, \text{ else } 0\,. \tag{147}$$

Using (50), the norm $\underset{r}{R}$ of $\underset{r}{\mathbf{y}}$ is found to be

$$\underset{r}{R} = \sqrt{m}/(g\sqrt{N^3})\,, \tag{148}$$

where $m>0$ is the cardinality of the index set A.

For any signal $\mathbf{y}$, the *spectrum* ϕ_α is defined as the normalized Hadamard

components,

$$\phi_\alpha = y_\alpha / R \,, \tag{149}$$

where R is the Euclidean norm of **y** given by (50); one has

$$\phi^\alpha \phi_\alpha = 1 \,. \tag{150}$$

The unstable stationary points $\underset{r}{y}$ of (147) have the spectrum

$$\text{if } \alpha \in A, \ \underset{r}{\phi}_\alpha = 1/\sqrt{m} \text{ , else } 0 \,. \tag{151}$$

The spectra given by (151) play an important role in the theory; we call them *principal spectra*. The integer m is called the *order* of the principal spectrum.

The principal spectra of order one have a unit vector for ϕ_α. We call such spectra *pure*; they have signals that are proportional to a Hadamard vector. The principal spectra of order two have two components equal to $1/\sqrt{2}$ and all other components zero; they have signals **y** that are proportional to the sum of two Hadamard vectors. The number of principal spectra of order m is $C(N,m) = \frac{N!}{m!\,(N-m)!}$. The total number principal spectra is $2^N - 1$, which is nearly as large as the number of corner points of the solid hypercube.

The set of unstable stationary points $\underset{r}{y}$ given by (147) is called the *ridge set*, since they lie on the ridge of the central crater. For ease of reference we state the result

Theorem 8: The ridge set has principal spectra.

The spectrum of **y** may be seen as a unit vector along **y** referred to the Hadamard base. It is of interest to write (149) as

$$y_\alpha = R\, \phi_\alpha \,, \tag{152}$$

and express the energy in terms of R and ϕ_α. For the unsubtracted dynamics the energy (133) then takes the form

$$E(R,\phi_\alpha) = -\tfrac{1}{3} \sqrt{N^3}\, R^3\, G + R^2/(2g) \,, \tag{153}$$

where

$$G = \sum_\alpha \phi_\alpha{}^3 \,. \tag{154}$$

Expression (153) for the energy can be used for the further exploration of the energy landscape, in two ways. One way is to fix the ϕ_α and consider E as function of R. This amounts to seeing how the energy changes along a ray through the origin. The direction of the ray is set by the spectrum ϕ_α. In the second method the radius R is fixed, and we regard E as function of the ϕ_α. This function shows how the energy varies over the hypersphere H_R centered at the origin. Together these two cuts provide a complete picture of the salient features of the energy landscape in the proportional region. We proceed with

the first cut.

For fixed G, the function E(R) of (153) has a minimum at R=0, and a maximum at

$$R = \underset{m}{R} = 1/(g\sqrt{N^3}\, G) . \tag{155}$$

The R=0 minimum is, of course, the stable point at the origin found before. The energy at the maximum is

$$E(\underset{m}{R}) = 1/(6g^3\, N^3\, G^2) . \tag{156}$$

E(R) is zero at $R = \frac{3}{2}\underset{m}{R}$. It follows that there is a central crater surrounded by a ridge; farther out the energy decreases to negative values. The ridge height given by (156) depends on the value G, which by (154) depends on the direction of the ray (152). Hence, the ridge is undulated, so that it has passes and peaks. The stationary points on the ridge have principal spectra, by Theorem 8. For the principal spectra we have

$$G = 1/\sqrt{m} . \tag{157}$$

Using (157) in (155) and (156) gives, for principal spectra,

$$\underset{m}{R} = \sqrt{m}/(g\sqrt{N^3}) , \tag{158}$$

and

$$E(\underset{m}{R}) = m/(6g^2N^3) , \tag{159}$$

which shows that the ridge can be passed easiest for m=1, i.e., along a ray pointing to a Hadamard corner. Also, among the principal spectra, the pass is located nearest the origin along those directions.

We need to be concerned that the points of the ridge set lie indeed in the proportional region. For the piecewise linear output function, this region is given by

$$|y_a| < 1 . \tag{160}$$

For the ridge set we have, from (147) and a Hadamard transform,

$$y_a = \frac{1}{gN^2} \sum_{\alpha \in A} h_{\alpha a} . \tag{161}$$

Since the sum over m Hadamard vectors has components which have at most the magnitude m, a sufficient condition for the signal **y** of (161) to lie in the proportional region is

$$g > m/N^2 . \tag{162}$$

This condition is satified for any $0<m\leq N$ if

$$g>1/N . \tag{163}$$

For the hyperbolic tangent output function, the proportional region is smaller than that given by (160) by a factor that depends on the approximation accuracy required. By (C9), the relative accuracy of the linear approximation is about $y^2/3$, where y is the

maximum magnitude among the components y_a. For instance, a 4% accurate linear approximation results if $|y_a|<1/3$, $\forall$ a. In order to obtain such accuracy for the ridge set it would be sufficient to replace (163) by

$$g>3/N\ . \tag{164}$$

Conditions (163) and (164) are easily satisfied, because in practice we want large gains in order that the DLS be fast.

In practice, the central crater is very small. For example, for N=16 and g=50, the radius R_m ranges from 1/3200 to 1/800, as m ranges from 1 to 16. These radii should be compared to the radius R=16 for points of the I_{16}.

Next, we take the hyperspherical cross section. On any hypersphere H_R of points y_α,

$$y^\alpha y_\alpha = R^2, \tag{165}$$

which lie in the proportional region, the stationary points of E given by (153), for fixed R, are the stationary points of G, subject to the subsidiary condition (150). Those points can be determined by using a Lagrangian multiplier,

$$0 = \partial_\alpha (G + \lambda (\phi^\alpha \phi_\alpha - 1)) = 3\phi_\alpha{}^2 + 2\lambda\phi_\alpha\ . \tag{166}$$

The multiplier λ is calculated by multiplying with ϕ^α and summing over α; the result is

$$\lambda = -\frac{3}{2} G\ . \tag{167}$$

Substitution in (166) gives

$$0 = \phi_\alpha(\phi_\alpha - G)\ , \tag{168}$$

and it follows that for a stationary point of the energy on the hypersphere H_R there is an index set A, such that

$$\text{if } \alpha \in A,\ \phi_\alpha = G,\ \text{ else } 0. \tag{169}$$

The value of G can be determined from (154):

$$G = m\ G^3\ , \tag{170}$$

where m is the cardinality of the set A. The solutions of (170) are G=0 and $G = \pm 1/\sqrt{m}$. It follows that the stationary points of the energy on the hypersphere H_R have the spectra

$$\begin{gathered} \text{if } \alpha \in A,\ \phi_\alpha = 1/\sqrt{m},\ \text{else } 0, \\ \text{or} \\ \text{if } \alpha \in A,\ \phi_\alpha = -1/\sqrt{m},\ \text{else } 0\ , \end{gathered} \tag{171}$$

where A is any index set, and m is its cardinality. From (171) and (151) we have

Theorem 9: For a DLS with r=0, and an output function that is either a hyperbolic tangent or a piecewise linear function, the stationary points of the energy function constraint to a

hypersphere H_R have, in the proportional region, ±principal spectra.

It can be shown that for $1<m<N$ these stationary points are saddle points. For $m=1$, i.e., the pure spectra, the positive solutions (171) are energy minima, and the negative solutions are energy maxima. For $m=N$, the positive solution (171) is an energy maximum, and the negative solution is an energy minimum.

The exploration of the energy landscape so far was restricted to the proportional region. Farther out in signal space, near the boundaries of the solid hypercube $J_N=[-1,1]^N$, the energy surface has a lip which turns up near these boundaries, and provides containment of the state point. In the equations of motion the lip corresponds to the term $-v_a$ outside the proportional region. This term can assume any value in balancing the equations of motion at equilibrium. Large magnitudes of v_a correspond to the large slopes available on the lip of the energy surface. The lip structure is expressed by the features of the function Ω that are not described by the approximation $R^2/(2g)$ valid in the proportional region.

DECOMPOSITION INTO LONGITUDINAL AND TRANSVERSE DYNAMICS

It is useful to introduce the notions of longitudinal and transverse parts of the equations of motion. This involves defining the *longitudinal* part $\mathbf{p}_{\|}$ of any vector $\mathbf{p}$ in signal space as the part of $\mathbf{p}$ in the direction of $\mathbf{y}$, and the *transverse* part $\mathbf{p}_{\perp}$ as the part of $\mathbf{p}$ perpendicular to $\mathbf{y}$. One has

$$\mathbf{p}_{\|} = \frac{\mathbf{p}\cdot\mathbf{y}}{|\mathbf{y}|}\frac{\mathbf{y}}{|\mathbf{y}|}, \quad \mathbf{p}_{\perp} = \mathbf{p} - \mathbf{p}_{\|} . \tag{172}$$

Application to the activation velocity vector $\dot{\mathbf{v}}$ defined by the equations of motion (34) for unsubtracted dynamics gives the *longitudinal* equation of motion

$$\dot{v}_{\|\alpha} = -v_{\|\alpha} + R\sqrt{N^3}Gy_\alpha + r_{\|\alpha}, \tag{173}$$

and the *transverse* equations of motion

$$\dot{v}_{\perp\alpha} = -v_{\perp\alpha} + \sqrt{N^3}(y_\alpha^2 - RGy_\alpha) + r_{\perp\alpha}, \tag{174}$$

where G is given by (154).

For ease of discussion we write the transverse equations of motion as

$$\dot{v}_{\perp\alpha} = -v_{\perp\alpha} + F_\alpha, \tag{175}$$

where

$$F_\alpha = \sqrt{N^3}(y_\alpha^2 - RGy_\alpha) + r_{\perp\alpha} . \tag{176}$$

ADIABATIC FAKE DYNAMICS

The split of the equations of motion into longitudinal and transverse parts may be used in the following manner. For fixed R, the transverse equation (175) describes the activation velocity on a centered hypersphere H_R of radius R in signal space. Suppose we constrain the signal state point to remain on H_R, while allowing the transverse dynamics given by (175). The state point on H_R will move to the *transverse equilibrium*, where $\dot{\mathbf{v}}_\perp = 0$, and

$$\mathbf{v}_\perp = \mathbf{F}\,, \tag{177}$$

by (175). We delete the longitudinal equation of motion (173) from the dynamics for now. In the modified dynamics, we fix R, and wait for transverse equilibrium to be reached on H_R. After that has happened, we move to the "next" sphere H_{R+dR}, and wait for transverse equilibrium to be reached on that sphere. Then, we proceed to the next hypersphere, etc. We call this fake dynamics *transverse adiabatic dynamics.* This dynamics may be executed either forward or backward, as the radius R of the hypersphere is increased or decreased in succession. The solution $\mathbf{y}$ of (177) depends on the parameter R; the path $\mathbf{y}(R)$, R = 0 to R_m, is called a *transverse adiabatic path.* R_m is the maximum value of R for which $\mathbf{y}$ lies in the solid hypercube J_N.

Let the signal point $\mathbf{y}_s$ be stationary in the true dynamics. Choose R such that $\mathbf{y}_s$ lies on H_R. Then, move backward through the transverse adiabatic path. As the radius R is diminished, the signal eventually falls in the proportional region, where

$$y_a = gv_a\,. \tag{178}$$

But this implies that $\mathbf{v}$ is longitudinal, i.e.,

$$\mathbf{v}_\perp = 0. \tag{179}$$

With (177) it follows that in the proportional region we have

$$\mathbf{F} = 0\,. \tag{180}$$

This shows

Theorem 10: For a DLS with unsubtracted dynamics, let $\mathbf{y}_s$ be a stationary point of the true dynamics, and let P be the transverse adiabatic path through $\mathbf{y}_s$. Then, $\mathbf{F} = 0$ along P in the proportional region.

It follows that there is a correspondence between stationary points and the solutions of Eq. (180). We proceed to find these solutions, for the simple case $\mathbf{r} = 0$. Then (176) becomes

$$F_\alpha = \sqrt{N^3}(y_\alpha{}^2 - RGy_\alpha)\,. \tag{181}$$

With (152) this may be written

$$F_\alpha = R^2\sqrt{N^3}(\phi_\alpha{}^2 - G\phi_\alpha)\,, \tag{182}$$

and (180) gives

$$\phi_\alpha{}^2 - G\phi_\alpha = 0. \tag{183}$$

But this is the same as condition (168), so that the transverse equilibria on H_R coincide with the stationary points of the energy function constraint to the hypersphere H_R. Of course, such coincidence was expected. With Theorem 9 we have the result

Theorem 11: For a DLS with unsubtracted dynamics and $r=0$, the transverse equilibria in the proportional region have ±principal spectra.

Each of the solutions of (180) corresponds to a stationary point of the true dynamics, by Theorem 10. A stationary point $\mathbf{y}_s$ corresponding to the solution $\mathbf{y}_o$ of (180) must lie on the transverse adiabatic path $\mathbf{y}(R)$ through $\mathbf{y}_o$. Its location on the path is such that there is longitudinal equilibrium, as stated by (173) with $\dot{\mathbf{v}}_{\|} = 0$,

$$v_{\|\alpha} = R\sqrt{N^3}Gy_\alpha\,; \tag{184}$$

r has been set to zero, as before.

For $r=0$, signals with principal spectra of orders one and two have transverse adiabatic paths that are straight lines through the origin, as will be discussed in the next section. For these cases, the location of the stationary point $\mathbf{y}_s$ on the transverse adiabatic path (180) may be determined from the longitudinal equilibrium condition (184). Since $\mathbf{v}_\perp = 0$ along the radial transverse adiabatic path, (184) gives

$$\mathbf{v}_s = R\sqrt{N^3}G\mathbf{y}_s\,. \tag{185}$$

For a principal spectrum of order m, we have

$$G = \pm 1/\sqrt{m}\,, \tag{186}$$

and (185) becomes

$$\mathbf{v}_s = \pm\frac{R\sqrt{N^3}}{\sqrt{m}}\mathbf{y}_s\,. \tag{187}$$

A solution of (185) may be obtained by writing

$$\mathbf{v}_s = v\mathbf{e}\,,\ \mathbf{y}_s = R\mathbf{e}\,; \tag{188}$$

where $\mathbf{e}$ is the unit vector along $\mathbf{y}$. Eq. (187) then gives

$$v = R^2\sqrt{N^3}G\,, \tag{189}$$

with G=1 for m=1, and G=$1/\sqrt{2}$ for m=2. For the hyperbolic tangent output function (18) the function v(R) implied by R=s(v) is

$$v(R)=\frac{1}{2g}\ln\frac{1+R}{1-R}, \tag{190}$$

by (C2), so that the solutions of (187) must satisfy the equations

$$m=1: \qquad \frac{1}{2g}\ln\frac{1+R}{1-R}=R^2\sqrt{N^3}, \tag{191}$$

$$m=2: \qquad \frac{1}{2g}\ln\frac{1+R}{1-R}=R^2\sqrt{(N^3/2)}. \tag{192}$$

The stationary points of the true dynamics for signals with principal spectra of orders one and two can be found by solving Eqs. (191) and (192) for R.

CONSERVATION OF PRINCIPAL SPECTRA

We have seen that, for unsubtracted dynamics and r=0, the transverse equilibria in the proportional region have ±principal spectra. This means that the adiabatic fake dynamics in the proportional region conserves ±principal spectra. Will this also happen in the true dynamics? Let ϕ_α be a principal spectrum of order m. Then,

$$\phi_\alpha{}^2-\phi_\alpha/\sqrt{m}=0 \tag{193}$$

as follows from (151). The importance of (193) is that it provides the possibility of replacing the quadratic term in the equations of motion by a linear term. For unsubtracted dynamics the equations of motion (34) then may, with (152), be cast in the form

$$\dot{v}_a=-v_a+R^2\sqrt{(N^3/m)}\,\phi_a+r_a, \tag{194}$$

where a Hadamard transform has been used to write ϕ_α in terms of ϕ_a. Suppose **r** has the same spectrum as **y**; then we may write

$$r_a=c\phi_a, \tag{195}$$

where c>0 is some fixed coefficient, and (194) becomes

$$\dot{v}_a=-v_a+R^2\sqrt{(N^3/m)}\,\phi_a+c\phi_a. \tag{196}$$

In the proportional region we have

$$v_a=\frac{1}{g}y_a=\frac{R}{g}\phi_a, \tag{197}$$

so that (196) may be written

$$\frac{d}{dt}(R\phi_a)=(-R+gR^2\sqrt{(N^3m)}+gc)\phi_a. \tag{198}$$

This equation is satisfied if

$$\dot{R}=-R+gR^2\sqrt{(N^3/m)}+gc\ , \tag{199}$$

and

$$\dot{\phi}_a=0\ . \tag{200}$$

consistent with (193). Hence we have as result

Theorem 12 Let ϕ be a principal spectrum. For a DLS with unsubtracted dynamics and $r=c\phi$, the principal spectrum ϕ is conserved in the proportional region.

Initial conditions and c can be chosen such that the function R(t) subject to (199) increases monotonically. Since the spectrum is constant as the signal point traverses the proportional region, the path is a straight line through the origin. The path is a staight line only if the spectrum is ±principal. Upon leaving the proportional region, the path generally curves away from a straight line, because of distortions in the spectrum produced by the nonlinearity of the output function. We know of two exceptions to this behavior, viz., the principal spectra with m=1 and m=2.

In the former case, the signal is proportional to a single Hadamard vector, and all ϕ_a (the spectrum components in the neuron frame) have the same magnitude, $1/\sqrt{N}$. The nonlinearity of the output function s(.) then applies uniformly to all components, so that in (196) the $|v_a|$ remain equal to each other. It follows that the signal path remains straight all the way to the stationary point in the Hadamard corner.

For the other case, m=2, the spectrum is proportional to the sum of two Hadamard vectors, say $\mathbf{h}_\beta$ and $\mathbf{h}_\gamma$,

$$\phi_a=(1/\sqrt{(2N)})(h_{\beta a}+h_{\gamma a})\ , \tag{210}$$

so that the components ϕ_a are $2/\sqrt{(2N)}$, $-2/\sqrt{(2N)}$, or 0. A zero value for ϕ_a gives $v_a=0$ by the antisymmetry of the neuron output function s(.), and that is consistent with (196). The other two possible values of ϕ_a lie symmetric with respect to 0, and will give equal magnitudes for the corresponding components of v_a, again by the antisymmetry of s(.); this is consistent with the equations of motion (196). It follows that the signal path remains straight outside the proportional region, all the way up to the stationary state. For sufficient gain, the stationary state will have components 0 and $\pm\lambda$, where λ is close to or equal to unity, depending on the output function used. By Theorem 3, such a stationary state is unstable.

SPURIOUS STATES SHIELD SPOILS EARLY DYNAMICS

The forgoing explorations provide an orientation and preparation for the main dynamics problem of the DLS: With the initial activation reset to zero, and the BLT output **u** presented to the DLS as external coupling, at what state will the DLS settle? We want this final state to be the Hadamard vector that is dominant in **u**.

In this regard we need to be concerned about the term $(N^2-4N)h_{1a}$ in the expression (66) for r_a. This term has the large magnitude N^2-4N, and it was added to r_a to shield against spurious states, as can be seen from the proof of Theorem 2. The term tilts the energy surface by a large amount in the direction of the first Hadamard vector, $\mathbf{h}_1$. As a result, the signal state point **y** will, in its gradient descent down the energy surface (in the proportional region) slide over the side of the gully that runs towards the dominant Hadamard point, and end up in either the adjacent gully, or in a gully several Hadamard vectors over. In any case, the state point will have left the gully which leads to the correct Hadamard point. This causes the DLS to settle at the wrong Hadamard point. The unwanted effect is largest close to the origin, where the undulating features in the energy landscape are subtle, so that the tilt of the energy surface has a large effect. Hence, the large threshold term $(N^2-4N)h_{1a}$, which was deployed as as shield against spurious stable states, spoils the early dynamics.

What is to be done? Either we have the protection against spurious states, and the wrong early dynamics, or we have the correct early dynamics, but face, later in the state development, the hazard of ending up at a spurious state. Under the circumstances, we choose the latter. The troublesome term $(N^2-4N)h_{1a}$ in the vector r_a is dropped, and we use the expression (41) or (44) for r_a in the external coupling or the initial value coupling. In either scheme, the vector r_a still contains a term that is proportional to δ_{a1}, with magnitude cN in the external coupling, and with magnitude c in the initial value coupling. These terms cause no problems with the dynamics, because they provide the same force in the direction of all Hadamard points.

We proceed with the investigation of the dynamics with this arrangement.

DYNAMICS IN THE PROPORTIONAL REGION

For unsubtracted dynamics, the Hadamard transform of the equations of motion in the proportional region is given by (48). These equations are uncoupled and they can be integrated as follows. The index α is temporarily suppressed. (48) may be written

$$dy/(y-y_+) - dy/(y-y_-) = (y_+-y_-)\, g\sqrt{N^3}\, dt\,, \tag{211}$$

where y_+ and y_- are the roots of the quadratic form

$$y^2-y/(g\sqrt{N^3})+r/\sqrt{N^3}. \tag{212}$$

We recall that for $r = 0$ the energy landscape features a central crater surrounded by an undulated ridge. Applying $r_a=\mu(u_a+c\delta_{a1})$ in the external couling scheme means tilting the energy surface. For very small μ this shifts the stable stationary point away from the origin. A second stationary point occurs at larger radius, and this point is unstable. Choosing progressively larger values for the coupling constant μ makes the two stationary points come closer, coalesce, and disappear. This corresponds to the discriminant of the

quadratic form (212) becoming smaller, zero, and negative. There is no stationary point in the proportional region iff there is an index α for which the quadratic form has complex roots; there is then no point at which $\dot{\mathbf{y}} = 0$. The three cases: different real roots, coinciding roots, and complex roots have different dynamics and must be considered seperately. Starting with the case that the roots $y_{\pm}$ are complex, we have

$$y_{\pm} = p\, e^{\pm i\eta}\,, \tag{213}$$

where

$$p = \sqrt{r}/N^{3/4}\,, \tag{214}$$

and

$$\cos\eta = 1/((2g\sqrt{r}\,N^{3/4})\,. \tag{215}$$

(211) may be rewritten as

$$dy/(y-y_{+}) - dy/(y-y_{-}) = 2i\nu dt, \tag{216}$$

where

$$\nu = \sqrt{(g^2 r\sqrt{N^3}-1/4)}\,. \tag{217}$$

The integral of (216) with initial condition $y(0)=0$ is

$$\ln\left(\frac{y - y_{+}}{y - y_{-}}\,\frac{y_{-}}{y_{+}}\right) = 2i\nu t\,, \tag{218}$$

which may be written

$$\frac{y - y_{+}}{y - y_{-}} = \frac{y_{+}}{y_{-}}\, e^{2i\nu t}\,. \tag{219}$$

This gives

$$y = y_{+}y_{-}\,\frac{1 - e^{2i\nu t}}{y_{-} - y_{+}e^{2i\nu t}} = y_{+}y_{-}\,\frac{e^{i\nu t} - e^{-i\nu t}}{y_{+}e^{i\nu t} - y_{-}e^{-i\nu t}} = y_{+}y_{-}\,\frac{e^{i\nu t} - e^{-i\nu t}}{e^{i(\eta+\nu t)} - e^{-i(\eta+\nu t)}}\,,$$

with the result

$$y_{\alpha} = p_{\alpha}\,\frac{\sin\ \nu_{\alpha} t}{\sin(\eta_{\alpha}+\nu_{\alpha} t)}\,, \tag{220}$$

where the index α has been reinstalled. With the index α shown, (214), (215), and (217) read

$$p_{\alpha} = \sqrt{r}_{\alpha}/N^{3/4}\,, \tag{221}$$

$$\cos\eta_{\alpha} = 1/(2g\sqrt{r}_{\alpha}\,N^{3/4})\,, \tag{222}$$

and

$$\nu_{\alpha} = \sqrt{(g^2 r_{\alpha}\sqrt{N^3} - 1/4)}\,. \tag{223}$$

In the external coupling scheme we have

$$r_{\alpha} = \mu(u_{\alpha} + c\sqrt{N}h_{\alpha 1})\,, \tag{224}$$

where $\mathbf{u}$ is the BLT output (3). In (224) we have $h_{\alpha 1}=1\,,\ \forall\ \alpha$. In the interest of readability, we will drop the $h_{\alpha 1}$ and write (224) as

$$r_\alpha = \mu(u_\alpha + c\sqrt{N}) , \tag{225}$$

committing a notational sin in tensor calculus since in (225), a scalar $c\sqrt{N}$ appears to be added to a vector u_α. Using the Hadamard transform of (3) in (225) gives

$$r_\alpha = \mu\sqrt{N}(c_\alpha + c) , \tag{226}$$

where

$$c_\alpha = \mathbf{x}.\mathbf{q}_\alpha . \tag{227}$$

Recall that $\mathbf{x}$ is the SRM input vector, and the $\mathbf{q}_\alpha$ are the bipolar stored vectors. Use of (226) gives for (221), (222), and (223)

$$p_\alpha = \sqrt{(\mu(c_\alpha + c)/N)} , \tag{228}$$

$$\cos \eta_\alpha = 1/(2gN\sqrt{(\mu(c_\alpha + c))} , \tag{229}$$

$$\nu_\alpha = \sqrt{(g^2N^2\mu(c_\alpha + c) - 1/4)} . \tag{230}$$

So much for the case that the quadratic form (212) has complex roots. Two other cases remain to be investigated: the two roots coinciding, and two different real roots.

For the coinciding roots, the discriminant

$$D_\alpha = 1/(g^2N^3) - 4r_\alpha/\sqrt{N^3} \tag{231}$$

is zero; this implies

$$\mu(c_\alpha + c) = 1/(4g^2N^2) . \tag{232}$$

The single root is

$$y_\infty = 1/(2g\sqrt{N^3}) , \tag{233}$$

and the equation of motion (48) with initial condition $y_\alpha(0)=0$ has the integral

$$y_\alpha = y_\infty \left(1 - \frac{1}{1+y_\infty\ g\ \sqrt{N^3}t}\right) . \tag{234}$$

This shows that

$$t \geq 0, \quad 0 \leq y_\alpha < y_\infty = 1/(2g\sqrt{N^3}) . \tag{235}$$

For large dimension and gain, y_∞ is small.

Next, we consider the case with two different real roots. There are three subcases: $c_\alpha + c < 0$, $c_\alpha + c = 0$, and $c_\alpha + c > 0$. For $c_\alpha + c = 0$, the equation of motion (48) with initial condition $y_\alpha(0)=0$ gives $y_\alpha = 0$ for all times. For $c_\alpha + c < 0$, the roots are written

$$y_+ = p\, e^{\eta'}, \; y_- = -p\, e^{-\eta'} , \tag{236}$$

where

$$p = \sqrt{(-r)}/N^{3/4} , \tag{237}$$

and η' is determined by

$$\sinh \eta' = 1/(2gN^{3/4}\sqrt{(-r)} . \tag{238}$$

With

$$\nu' = \sqrt{(1/4 - r g^2 \sqrt{N^3})} , \tag{239}$$

the equation of motion (48) with initial condition y(0)=0 has the integral

$$y = -p \frac{\sinh \nu' t}{\cosh(\eta' + \nu' t)} . \tag{240}$$

For large t we have

$$y \to -p\, e^{-\eta'} = y_- . \tag{241}$$

Since the function y(t) of (240) is monotone for $t \geq 0$, we have

$$t \geq 0, \qquad 0 \leq y_\alpha < y_{\alpha-} , \tag{242}$$

where the index α has been reinstalled. Use of (226) gives

$$y_{\alpha-} = (1/2 - \sqrt{(1/4 - \mu(c_\alpha + c) g^2 N^2)})/(g\sqrt{N^3}) . \tag{243}$$

For

$$-\mu(c_\alpha + c) g^2 N^2 << 1 \tag{244}$$

we have

$$y_{\alpha-} \simeq \mu(c_\alpha + c) g \sqrt{N} , \tag{245}$$

which is small for large dimension and gain, considering (244). For $\mu(c_\alpha + c) g^2 N^2$ of order unity, $y_{\alpha-}$ of (243) is small for large g and N. For $-\mu(c_\alpha + c) g^2 N^2 >> 1$ we have

$$|y_{\alpha-}| \simeq \sqrt{(-\mu(c_\alpha + c)/N)} , \tag{246}$$

which is small for large dimension. By (242) it follows that $|y_\alpha|$ always remains small if g and N are large.

Finally, for the case with real roots $y_\pm$ and $c_\alpha + c > 0$ we have

$$y_\pm = p\, e^{\pm \eta'} , \tag{247}$$

$$p = \sqrt{r}/N^{3/4} , \tag{248}$$

$$\cosh \eta' = 1/(2g\sqrt{r}\, N^{3/4}) , \tag{249}$$

$$\nu' = \sqrt{(1/4 - r g^2 N^{3/2})} . \tag{250}$$

The integral of the equation of motion (48) with initial condition y(0)=0 is

$$y = p \frac{\sinh \nu' t}{\sinh(\eta' + \nu' t)} . \tag{251}$$

For large t we have

$$y \to p\, e^{-\eta} = y_{-}. \tag{252}$$

Since the function y(t) given by (251) is monotone, and y(0)=0, it follows that

$$t \geq 0, \quad 0 \leq y_{\alpha} < y_{\alpha -} = (1-\sqrt{(1-4\mu(c_{\alpha}+c)g^2N^2)}/(2g\sqrt{N^3}) , \tag{253}$$

where the index α has been restored. Since the case considered has two different real roots and $c_{\alpha}+c>0$, we have $0<4\mu(c_{\alpha}+c)g^2N^2<1$ in (253), and it follows that

$$0<1-\sqrt{(1-4\mu(c_{\alpha}+c)g^2N^2)}<1 . \tag{254}$$

Hence, $y_{\alpha -}$ is small for large g and N.

We have seen that, for the case that the quadratic form (212) has real roots, $|y_{\alpha}|$ has a bound that is small for large g and N. Hence, for a DLS with unsubtracted dynamics, large gain and dimension, using the external coupling scheme, the signal **y** can get substantially away from the origin only if there exists an index α for which the roots of the quadratic form (212) are complex, i.e., if

$$\exists\ \alpha \text{ such that } g^2N^2\mu(c_{\alpha}+c) > 1/4 ; \tag{255}$$

for all such α, the Hadamard component y_{α} of the DLS signal **y** is given by (220). We consider here only cases such that the dominant Hadamard vector in the BLT output **u** is unique, i.e.,

$$c_{\beta} > c_{\alpha} , \ \forall\ \alpha \neq \beta . \tag{256}$$

The index β for which (256) is true is called the *dominant index*. Eq. (227) shows that if two coefficients c_{γ} are different, they differ by at least 2. Since $c_{\alpha} \epsilon [-N,N]$ because of the bipolar nature of $\mathbf{q}_{\alpha}$ and **x** and (227), (256) implies that

$$c_{\beta} \epsilon [-N+2,N] . \tag{257}$$

It follows that the condition (255) written for the dominant index β,

$$g^2N^2\mu(c_{\beta}+c) > 1/4 \tag{258}$$

is satisfied if we choose

$$c=N , \tag{259}$$

and

$$\mu > 1/(8g^2N^2) . \tag{260}$$

For our DLS to work, the signal **y** must traverse the proportional region. With the results of this section this means that in the proportional region and for large dimension and gain, at least one Hadamard component must grow according to (220). Clearly, one of these

Hadamard components is the dominant Hadamard component y_β. It follows that for the dominant index β the roots $y_\pm$ must be complex; this is the case if (260) is satisfied. Hence we have

Lemma 5: For a DLS with unsubtracted dynamics, externally coupled to a BLT, with large dimension and gain, and with the coupling constant satisfying (260), the dominant Hadamard component grows in the proportional region according to (220).

So far in this section we have considered the dynamics in the proportional region for a DLS with external coupling to the BLT output **u**. We briefly consider the other coupling, i.e., initial value coupling. We now have expression (44) for the vector **r**. The Hadamard components are

$$r_\alpha = (c/\sqrt{N})h_{\alpha 1} . \tag{261}$$

For complex roots $y_\pm$ we have from (216), and the initial value

$$y_\alpha(0)=\mu g u_\alpha \tag{262}$$

the solution

$$y_\alpha = p\,\frac{p\,\sin\nu t\; + g\mu u_\alpha \sin(\eta-\nu t)}{p\,\sin(\eta+\nu t)\; - g\mu u_\alpha \sin\,\nu t} , \tag{263}$$

where

$$p=\sqrt{c}/N , \tag{264}$$

$$\cos\,\eta = \frac{1}{2g\sqrt{(cN)}} , \tag{265}$$

and

$$\nu=\sqrt{(g^2cN-1/4)} . \tag{266}$$

The condition for the roots to be complex is

$$c>1/(4Ng^2) . \tag{267}$$

The roots must be complex in order that the signal **y** can traverse the proportional region. Condition (267) is satisfied if we choose

$$c=1/(2Ng^2) . \tag{268}$$

For this value of c, (264) to (266) become

$$p=1/(g\sqrt{(2N^3)}) , \tag{269}$$

$$\cos\,\eta = 1/\sqrt{2} , \tag{270}$$

$$\nu = 1/2 \tag{271}$$

(270) shows that $\eta=\pi/4$, and (263) becomes

$$y_\alpha = p\,\frac{(\sqrt{2}-z_\alpha)\ \sin\ \nu t\ +\ z_\alpha \cos\ \nu t}{\cos\ \nu t\ -(z_\alpha\sqrt{2}-1)\ \sin\ \nu t}\ , \tag{272}$$

where

$$z_\alpha = g^2 \mu u_\alpha \sqrt{(2N^3)}\ . \tag{273}$$

There is a time $t_{\alpha\infty}$ such that the function $y_\alpha(t)$ given by (272) is monotone and increasing in the interval

$$0 \le t < t_{\alpha\infty}\ , \tag{274}$$

and

$$y(t_{\alpha\infty}) = \infty\ . \tag{275}$$

t_∞ decreases for increasing z_α. Since z_α is largest for the dominant index β, t_∞ is smallest for y_β. It follows that y_β given by (272) increases indefinitely as t approaches $t_{\beta\infty}$. Of course, at sometime during this process the signal **y** leaves the proportional region. The singularity at $t_{\beta\infty}$ dominates the behavior of y_β as t gets close to $t_{\alpha\infty}$. This results in spectral purification, as will be discussed in the next section. We note here a disadvantage of the initial value coupling: if $u_\beta<0$, the initial signal has negative y_β, and it takes time for the y_β to become positive (as required, because we want the signal to go to $\mathbf{h}_\beta$) by the action of the threshold term **r** given by (261). The external coupling does not have this delay, since with c given by (259), $c_\beta+c$ in (226) is always positive, so that the y_β becomes positive immediately after reset of the DLS. In this regard it should also be noted that it takes time to reset the activation to any value, be it $\mu(u_a+Nc\delta_{a1})$ or zero, because amplifiers have a finite slew rate.

SPECTRAL PURIFICATION IN THE PROPORTIONAL REGION

A surprising property of the DLS is that already in the proportional region, long before the signal **y** comes close to the boundaries of the solid hypercube J_N , the spectrum gets purified towards the dominant Hadamard vector. This purification is due to the activation nonlinearity and the mathematical nature of Hadamard vectors. The spectral purification plays an important part in the development of the state, which, starting from the origin (by reset), makes its way towards the dominant Hadamard point $\mathbf{h}_\beta$ by gradient descent along the energy surface. Although spectral purification occurs also for initial value coupling, this will not be investigated here, because of the practical disadvantage of this coupling as noted above, and this report is getting too long. For the external coupling, the BLT output **u** is coupled to the DLS by means of the term

$$r_\alpha = \mu(u_\alpha + c\sqrt{N}) \tag{276}$$

in the Hadamard transform (48) of the equations of motion; unsubtracted dynamics is used here. The BLT output **u** is given by (3), with the Hadamard components

$$u_\alpha = \sqrt{N}\, c_\alpha , \tag{277}$$

where

$$c_\alpha = \mathbf{x}.\mathbf{q}_\alpha , \tag{278}$$

as given by (4). Remember that $\mathbf{q}_\alpha$ are the stored vectors, and $\mathbf{x}$ is the input vector. With (277) and the choice (259) for c, (276) gives

$$r_\alpha = \mu\sqrt{N}(c_\alpha + N) . \tag{279}$$

The dominant Hadamard vector in $\mathbf{u}$ is the Hadamard vector with the largest coefficient in the Hadamard expansion of $\mathbf{u}$. Hence, β is the dominant index in $\mathbf{u}$ iff

$$c_\beta > c_\alpha \ , \ \forall\ \alpha \neq \beta . \tag{280}$$

We can also consider the dominant index in the signal $\mathbf{y}$; it is the index β such that

$$y_\beta > y_\alpha , \forall\ \alpha \neq \beta . \tag{281}$$

Both (280) and (281) imply that the dominant index is unique; we restricted the inputs $\mathbf{x}$ such that this is the case. This condition excludes any input vector $\mathbf{x}$ which has more than a single nearest stored state.

Since for early times the signal $\mathbf{y}$ is about proportional to the vector $\mathbf{r}$, and $\mathbf{r}$ is given by (276), the dominant index in $\mathbf{u}$ is, for these times, the same as the dominant index in $\mathbf{y}$. Hence, for early times, i.e., close to the signal origin, these two definitions of dominant index may be interchanged.

Let β be the dominant index in $\mathbf{y}$, and let α the index with the next smaller Hadamard component c_α. We call the ratio

$$\mathbb{R} = y_\beta / y_\alpha \tag{282}$$

the *dominance ratio*. The dominance ratio changes through the proportional region. As the signal $\mathbf{y}$ traverses this region as described by the equation of motion (48), the norm of $\mathbf{y}$ increases, until the outer edge of the region is reached, say, at $y_\beta = K > 0$. The value of K depends on the other Hadamard components y_α, since the boundaries of the proportional region are given in terms of the y_a, not the y_α. It is convenient to have an estimate which does not equire knowledge of the other Hadamard components y_α. For the case of a piecewise linear output function, such an estimate is given by

Lemma 6: For a DLS with piecewise linear neuron output function, let the signal $\mathbf{y}$ have a unique dominant Hadamard component $y_\beta = K$, and let the dominance ratio be $\mathbb{R}$. ==>

$$K \leq \sqrt{N}/(1 + \frac{N-1}{\mathbb{R}}) .$$

Proof: By the inverse Hadamard transform we have

$$y_a=(1/\sqrt{N})\, h_a{}^{\alpha} y_\alpha=(1/\sqrt{N})(h_{a\beta}y_\beta+ \sum_{\alpha\neq\beta} h_{a\alpha}y_\alpha) . \tag{283}$$

With $y_\beta=K$ and $|y_\alpha|\leq \mathbb{R}K$, $\forall\ \alpha\neq\beta$, one has from (283)

$$|y_a|\leq (1/\sqrt{N})(K+(N-1)K/\mathbb{R}) . \tag{284}$$

For the piecewise linear output function, the proportional region is given by

$$|y_a|<1 . \tag{285}$$

Then, (284) implies that **y** lies in the proportional region if

$$K<\sqrt{N}/(1+ \frac{N-1}{\mathbb{R}}) . \quad] \tag{286}$$

We are interested in the dominance ratio at the point with $y_\beta=K$. Different cases need to be considered. First, we look at the case with

$$c_\beta=-N+2 . \tag{287}$$

Since the dominant index β is assumed to be unique, and c_α has the range [–N,N], one must have

$$c_\alpha=-N\ ,\ \forall\ \alpha\neq\beta . \tag{288}$$

At this point we choose the coupling constant as

$$\mu=1/(4g^2N^2) ; \tag{289}$$

this choice satisfies (260), and with (259) and (287), condition (255) is satisfied for index β. It follows that for index β the roots $y_\pm$ are complex, so that the development of y_β in time is described by (220). But for the other indices, (288) and (259) show that $c_\alpha+c=0$, so that $y_\alpha=0$, $\forall\ t\geq 0$. It follows that in this case the dominance ratio $\mathbb{R}$ is infinite, $\forall\ t\geq 0$ and, of course, **y** in the proportional region. Hence, in this case the spectrum is pure from the start.

The next case considered has

$$c_\beta=-N+4 . \tag{290}$$

If the next smaller Hadamard coefficient is $c_\alpha=-N$, then the considerations for the previous case apply and we have $\mathbb{R}=\infty$, $\forall\ t\geq 0$, in the proportional region. If the next smaller coefficient after c_β is

$$c_\alpha=-N+2 , \tag{291}$$

then both y_β and y_α are given by a formula of the type (220), and we have, with $y_\beta=K$,

$$\mathbb{R} = K/y_\alpha=K\ \sin(\eta_\alpha+\nu_\alpha t)/(p_\alpha \sin(\nu_\alpha t) , \tag{292}$$

where p_α, η_α, and ν_α are given by (228) to (230), with c=N; hence

$$p_\alpha=\sqrt{(\mu(c_\alpha/N+1)}\ , \tag{293}$$

$$\cos\ \eta_\alpha=1/(2gN\sqrt{(\mu(c_\alpha+N))}\ , \tag{294}$$

and

$$\nu_\alpha=\sqrt{(g^2N^2\mu(c_\alpha+N)-1/4)}\ . \tag{295}$$

For the coupling constant given by (289) we have from (293) to(295)

$$p_\alpha=\frac{\sqrt{(c_\alpha/N+1)}}{2gN}\ , \tag{296}$$

$$\cos\eta_\alpha=1/\sqrt{(c_\alpha+N)}\ , \tag{297}$$

and

$$\nu_\alpha=(1/2)\sqrt{(c_\alpha+N-1)}\ . \tag{298}$$

As the gain g or the dimension N is increased indefinitely, p_α of (296) tends to zero. Then, the equation $K=y_\beta$, for y_β given by (220), viz.,

$$K = p_\beta \frac{\sin\ \nu_\beta t}{\sin(\eta_\beta+\nu_\beta t)} \tag{299}$$

can be satisfied only if the denominator goes to zero, i.e., if

$$\eta_\beta+\nu_\beta t \to \pi\ . \tag{300}$$

With (290), (297), and (298), the statement (300) gives

$$t\to 4\pi/(3\sqrt{3})\ . \tag{301}$$

Using this result, together with the p_α, η_α, and ν_α given by (296), (297), and (298) for the c_α of (291) gives for the dominance ratio (292), after a short calculation,

$$\mathbb{R}\simeq 1.38KgN^{3/2}\ , \tag{302}$$

if either g or N is large. Using Lemma 6 with the equal sign together with (302) gives the result

$$\text{large g or large N}\ ,\quad \mathbb{R}\simeq 1.38g\nu^2-N+1\ . \tag{303}$$

Hence in this case the dominance ratio at the signal point with $y_\beta=K$ can be made as large as desired by choosing either N or g large enough.

Next, we consider the case with

$$c_\beta = -N+A \ , \ A \geq 6 \ , \tag{304}$$

The smallest dominance ratio results if the next smaller coefficient c_α is

$$c_\alpha = c_\beta - 2 \ . \tag{305}$$

For fixed $A \geq 6$, η_β, η_α, ν_β, and ν_α do not depend on N or g. As g or N is increased indefinitely, p_α tends to zero, and we must have (300), as before. Statements (301), (302), and (303) go through with different numerical factors; the important point is that these factors do not depend on g or N. The cases considered exhaust all possibilities for smallest $\mathbb{R}$. Hence we have

Theorem 13: For a DLS with unsubtracted dynamics, externaly coupled to a BLT with $r_a = \mu(u_a + N^2 \delta_{a1})$, and a coupling constant $\mu = 1/(4g^2N^2)$ ==> the dominance ratio $\mathbb{R}$ at the outer edge of the proportional region can be made arbitrarily large by taking either the gain g or the dimension N large enough.

FINAL PURIFICATION

After traversing the proportional region, the activation enters the region where the nonlinearity of the neuron output function is important. By Theorem 13, the dominance ratio at the edge of the proportional region can be made arbitrarily large by choosing a large gain or dimension. What happens to the dominance ratio in the region where the nonlinearity in the output function is important? We have

Theorem 14: For a DLS with unsubtracted dynamics and a piecewise linear neuron output function, externally coupled to the BLT output **u** by $r_a = \mu(u_a + N^2 \delta_{a1})$, with coupling constant $\mu = 1/(4g^2N^2)$, and with the activation reset to zero at the time of application of **u**, the signal **y** settles at the dominant Hadamard vector in **u**, if the gain g is large enough.

Proof: For unsubtracted dynamics we have the equations of motion (34):

$$\dot{v}_a = -v_a + N \sum_\alpha h_{\alpha a} y_\alpha^2 + \mu(h^\alpha{}_a c_\alpha + N^2 \delta_{a1}) \ , \tag{306}$$

where (41), (3), and (259) have been used. Showing the β term in the sum separately, this is written as

$$\dot{v}_a = -v_a + N \left(h_{\beta a} y_\beta^2 + \sum_{\neq \beta} h_{\alpha a} y_\alpha^2\right) + \mu(h^\alpha{}_a c_\alpha + N^2 \delta_{a1}) \ . \tag{307}$$

Similarly, we write

$$y_a = (1/\sqrt{N})\, h^\alpha{}_a y_\alpha = (1/\sqrt{N}) \left(h_{\beta a} y_\beta + \sum_{\alpha \neq \beta} h_{\alpha a} y_\alpha\right) . \tag{308}$$

There are separate arguments for $a=1$ and $a \neq 1$. We start with the case $a \neq 1$. (307) gives

$$a \neq 1 \ , \quad \dot{v}_a = -v_a + N \left(h_{\beta a} y_\beta^2 + \sum_{\neq \beta} h_{\alpha a} y_\alpha^2\right) + \mu h^\alpha{}_a c_\alpha \ . \tag{309}$$

With the choice (289) for the coupling constant, one has for the last term

$$|\mu h^{\alpha}{}_{a} c_{\alpha}| \leq 1/(4g^2) , \tag{310}$$

since $|c_{\alpha}| \leq N$, by (278), and the bipolar nature of the vectors $\mathbf{x}$ and $\mathbf{q}_{\alpha}$. It follows that the $|\mu h^{\alpha}{}_{a} c_{\alpha}|$ can be made arbitrarily small by choosing the gain large enough. Next, consider the term $\sum_{\alpha \neq \beta} h_{\alpha a} y_{\alpha}^{2}$ in (309). Let $\mathbf{y}$ be at the edge of the proportional region. Then, $y_{\beta}=K$, and $|y_{\alpha}|<K/\mathbb{R}$, $\forall\ \alpha \neq \beta$, and it follows that

$$|\sum_{\alpha \neq \beta} h_{\alpha a} y_{\alpha}^{2}| \leq (N-1)K^2/\mathbb{R}^2 . \tag{311}$$

By Theorem 13, $\mathbb{R}$ can be made arbitrarily large by choosing g large enough. It follows that the term $\sum_{\alpha \neq \beta} h_{\alpha a} y_{\alpha}^{2}$ tends to zero as g goes to infinity.

Since at the edge of the proportional region we must have

$$|v_{a}| \leq 1/g, \ \forall\ a, \tag{312}$$

the v_{a} also tends to zero as g goes to infinity.

It follows that, for fixed dimension N and a fixed activation $\mathbf{v}$ at the edge of the proportional region, we have

$$a \neq 1, \quad \dot{v}_{a} \rightarrow N\, h_{\beta a} y_{\beta}^{2} , \quad \text{as } g \rightarrow \infty . \tag{313}$$

We still need to consider a=1. The equation of motion (306) gives

$$a=1 , \quad \dot{v}_{1} = -v_{1} + N \sum_{\alpha} h_{\alpha 1} y_{\alpha}^{2} + (h^{\alpha}{}_{1} c_{\alpha} + N^2) . \tag{314}$$

For the last term one has

$$|h^{\alpha}{}_{1} c_{\alpha} + N^2)| = |\sum_{\alpha} c_{\alpha} + N^2| \geq 2 , \tag{315}$$

provided that the dominant index is unique; the proof is left to the reader. It follows that

$$v_{1} > 0 , \ \forall\ t > 0 , \tag{316}$$

and for the stationary point

$$v_{1} \geq N y^{\alpha} y_{\alpha} + 2 . \tag{317}$$

From (308) we have

$$|\sum_{\alpha \neq \beta} h_{\alpha a} y_{\alpha}| \leq (N-1)K/\mathbb{R} \rightarrow 0 \ \text{ as } \ g \rightarrow \infty , \tag{318}$$

because of Theorem 13. Hence,

$$y_a \to (1/\sqrt{N}) h_{\beta a} y_\beta = (K/\sqrt{N}) h_{\beta a} \text{ , as } g \to \infty \text{ .} \tag{319}$$

At the edge of the proportional region, $\exists$ b such that $|y_b|=1$. By (319) one has, for this index b,

$$1 = |y_b| = K/\sqrt{N} \text{ , as } g \to \infty \text{ ,} \tag{320}$$

so that

$$K \to \sqrt{N} \text{ , as } g \to \infty \text{ .} \tag{321}$$

Hence, we have from (319)

$$y_a \to h_{\beta a} \text{ , } \forall \text{ a, as } g \to \infty \text{ .} \tag{322}$$

(322) implies

$$|y_a| \to 1 \text{ , } \forall \text{ a, as } g \to \infty \text{ .} \tag{323}$$

and

$$y_\alpha = (1/\sqrt{N}) h_\alpha{}^a h_{\beta a} = \sqrt{N} \delta_{\alpha\beta} \text{ .} \tag{324}$$

For a piecewise linear output function, the signal **y** does not always describe the state unambiguously, and we then need to work with the activation **v**. From (324) we have

$$a \neq 1, \dot{v}_a \to N^2 h_{\beta a} \text{ , as } g \to \infty \text{ .} \tag{325}$$

For large g, (322) shows that the activations v_a for a≠1 all have about the same value, 1/g. By (313), the same the time derivatives $\dot{v}_a$, a≠1. all have the same magnitudes, and their signs are the same as the signs of y_a. It follows that shortly after the activation **v** leaves the proportional region, all components v_a , a≠1 are still about the same, and all exceed 1/g in magnitude. Hence, at that time t_1 we have

$$a \neq 1, \quad y_a = h_{\beta a} \text{ exactly .} \tag{326}$$

Eq. (313) remains true, so that for a≠1 the changes of v_a brought about by $\dot{v}_a$ of (313) always have the same sign as v_a, and $y_a = h_{\beta a}$ does not change at all (remember s(.) is the piecewise linear function). Hence, for a≠1 a stationary point is reached with

$$v_a \to N^2 h_{\beta a} \text{ , as } g \to \infty \text{ .} \tag{327}$$

For a=1, we have for the stationary state, by (317),

$$v_1 > N y^\alpha y_\alpha + 2 \to N y_\beta{}^2 + 2 = N^2 + 2 \text{ , as } g \to \infty \text{ ,} \tag{328}$$

by (324). Since, for large g, $N^2+2 > 1/g$, we have for the stationary point

$$y_1 = 1 \text{ , for large g.} \tag{329}$$

(327), (322), and (329) show that, for large g, the state moves to the stationary point

$$y_a = h_{\beta a}\ ,\ \forall\ a. \quad \rfloor \tag{330}$$

Two comments are in order. First, the condition $\mu = 1/(4g^2N^2)$ on the coupling constant can be relaxed to an inequality $1/(4g^2N^2) \leq \mu < c/(4g^2N^2)$, where $c>1$ is some number that can be determined with further effort. This relaxation of the coupling constant condition is important in practice, where we want a margin for structural parameters. We have not used an inequality for μ in order to keep the proof simple. The estimations used can be sharpened considerably; the sharper the estimates, the larger c can be.

The second comment is of a more serious nature. However much the estimates used in the proof of Theorem 14 are sharpened, the maximum coupling constant μ allowed is expected to be still small, because the basic factor $1/(4g^2N^2)$ is so small ($3.9\text{x}10^{-7}$ for N=16 and g=50). The practical problem with such small coupling constants is that they make the DLS exceedingly slow, because a very long time would be needed to develop the DLS activation to appreciable levels, after reset to zero at the time that the BLT output **u** is applied. In practice we need large gain and a coupling constant of the order of unity, for the sake of speed. Moreover, the important practical applications of the SRM are expected to have large dimension N. It is clear that for such fast SRMs Theorem 14 does not provide an assurance of perfect associative recall. Numerical computations have shown excellent performance for N=16 and coupling constants as large as 0.2, but this does not imply that such machines with much larger dimension would work for practical values of the coupling constant.

The basic problem is that the spectral purification processes captured by Theorems 13 and 14 do not describe the powerful purification that goes on in a DLS for signals close to the boundary of the solid hypercube. Such purification is observed clearly on the computer. A strong effort is needed to study this point and to provide a theorem that covers practical values of the coupling constant.

NUMERICAL COMPUTATIONS

Computations have been carried out to investigate the associative recall of SRMs of dimensions N=8 and 16. In these computations, a bipolar vector **x** is presented to the SRM front stage, and the output **x**′ of the SRM is compared with the stored vectors $\mathbf{q}_\alpha$, α=1 to N . The SRM has perfect associative recall if for every bipolar input vector **x** the output **x**′ is the stored vector $\mathbf{q}_\alpha$ nearest **x**, provided that the nearest stored vector is unique. The latter condition is accounted for by letting the computer skip input vectors **x** which have multiple nearest stored vectors. With these exceptions, the test of associative recall is applied to all vectors **x** of the N dimensional hypercube $I_N = \{-1,1\}^N$.

The N stored bipolar vectors $\mathbf{q}_\alpha$ were chosen at random, but were kept fixed during the run of **x** over the I_N. In the early stages of the project, before the theory was adequately developed, we had some notions that complelled us to store only N−1 vectors, which were subject to the condition that their first component always be +.

The SRM front stage is the BLT, which in the computations performs a linear

transformation of **x** by the matrix B, as shown in (1). The matrix B is the sum of outer products of the stored vectors and their labels, which are Hadamard vectors. The label for the stored vector $\mathbf{q}_\alpha$ is the Hadamard vector $\mathbf{h}_\alpha$. The BLT output is the vector **u**. In calculating **u**, the response time of the BLT amplifiers has been neglected. This was done to speed up the computations, and because a finite response time is not expected to change the results of the associative recall computations. The basis for this expectation is the absence of feedback in the BLT; there is no settling of **u** other than the smooth growth from zero to the full output.

The BLT output **u** is presented to the rear stage, the DLS, for further processing. In the computations, both coupling schemes, the initial value coupling, and the external coupling, have been used. In early stages of the project we used initial value coupling out of fear that external coupling would make spurious states possible. This is ideed the case, but we have since found out that, although spurious states exist, they are dynamically inaccessible in the external coupling scheme, provided that the coupling constant and gain are chosen properly. The initial value coupling was first applied with very small coupling constants, typically 10^{-7}. This coupling constant is just large enough to place the initial activation outside of the central crater, so that the state does not fall back to the stable origin. For N=16, perfect associative recall was found. However, the computations were very slow because of the small coupling constant. In practice, this would show up as a slow settling of the DLS. Moreover, in practice the weak initial value coupling would be vulnerable to noise. It was decided to diminish this vulnerability by increasing the coupling constant. It was found that then the gain must be diminished, in order to keep the recall perfect. The small gain was found to slow down the DLS . In response to these difficulties, we reexamined the external coupling, and started bearing down on the question of dynamic accessibility of the spurious states which are brought into existence by removing the safe threshold (the term $(N^2-4N)h_{1a}$ in (66)).

Extensive computations were done for the external coupling scheme. The strength of this coupling, i.e., the coupling constant μ in (41) affects the settling speed of the DLS, in a similar way as for the initial value coupling. In this regard, Theorem 14, which guarantees that the DLS works if the gain is chosen large enough, is subject to the condition (289) on μ, and the coupling constant values allowed are much too small in practice. Therefore, the numerical computations were done with much larger coupling constants, typically 0.2. Perfect associative recall was found for these SRMs, with gains of 50. We were interested in the spectrum purification, as the state developed from reset at zero to the dominant Hadamard corner. A very strong purification was noticed upon entering the partially supercritical region. The entering of this region was easily spotted when the piecewise linear output function was used, since then the enirely subcritical region is given by $|v_a|<1/g$, where g is the gain. The strong purification observed in the supercritical region is not exposed by Theorems 13 and 14, and hence, these theorems do not get at what makes the DLS work with practical values of the coupling constant.

Besides the two coupling schemes, a choice had to be made between subtracted and unsubtracted connection tensors, and between using y_1 clamping or not. The computations range over a number of cases, but do not cover all possibilities because of the computer time involved. The associative recall computations require, for each input vextor **x**, the numerical integration of N coupled nonlinear differential equations of first order. For N=16, the computation runs over 2^{16}=64K input vectors, although some of these are skipped because of multiple nearest vectors.

On a number of occasions it was observed that y_1 clamping speeds up the settling of

the DLS, and we expect this to be true in general.

The numerical integration of the DLS equations of motion was done in time steps which typically had the duration $\delta t=10^{-5}$. The time step must be compared to the RC time of unity for the equations of motion in the normalized form (15).

We need to discuss the cutoff for the numerical integration. After considerable experimentation, we settled on a scheme in which the integration is terminated if, for all indices b=1 to N, $\{|v_b|\geq 5/g$ and $|v_b(t)|\geq|v_b(t-10\delta t)|\}$. The factor 10 in $t-10\delta t$ was chosen to accomodate fluctuations in stochastic computations which use the same computer program. The stochastic computations will not be discussed here, but will be reported separately.

The following is a partial list of associative recall computations performed. All these runs showed perfect associative recall.

1) N=16, initial value coupling with $\mu=1x10^{-7}$, unsubtracted dynamics, y_1 clamping. Hyperbolic tangent output function with gain of 10.

2) N=16, initial value coupling with $\mu=1x10^{-7}$, subtracted dynamics, no y_1 clamping. Hyperbolic tangent output function with a gain of 10.

3) N=8, initial value coupling by Eq. (45) with $\mu=0.0001$ to 0.1, with zero threshold (c=0 in Eq. (44)). piecewise linear output function with gain of 0.25. Unsubtracted dynamics with y_1 clamping.

4) N=8, external coupling by Eq. (41) with c=0 and $\mu=1.0$. N unrestricted stored bipolar vectors. Unsubtracted dynamics without y_1 clamping. Piecewise linear output function with gain of 20.

5) N=16, external coupling by Eq. (41) with c=0 and $\mu=0.5$. N−1 stored bipolar vectors with first component +. Subtracted dynamics with y_1 clamping. Hyperbolic tangent output function with gain of 50.

6) N=16, external coupling by Eq. (41) with c=0, and $\mu=0.2$. N−1 stored bipolar vectors with first component +. Unsubtracted dynamics with y_1 clamping. Hyperbolic tangent output function with gain of 50.

7) N=16, external coupling by Eq. (41) with c=0, and $\mu=0.2$. N−1 stored bipolar vectors with first component +. Unsubtracted dynamics with y_1 clamping. piecewise linear output function with gain of 50.

8) N=16, external coupling by Eq. (41) with c=N, and $\mu=0.2$. N unrestricted bipolar stored vectors. piecewise linear output function with gain of 50. Unsubtracted dynamics. No y_1 clamping.

9) Same as 8), but with y_1 clamping.

CONCLUSION

The main result is contained in Theorem 14. This theorem assures perfect associative recall of an SRM consisting of a BLT, externally coupled to a DLS built as a quadratic Hadamard memory, if the coupling constant is as specified, and the gain is large enough. The theorem holds for any dimension that is a power of 2. The theorem is derived for unsubtracted dynamics and a piecewise linear output function. It is good to have such a theorem that is valid for the large dimensions that are expected to be important in applications. The condition of large gain in Theorem 14 is likely to be met in practice, because large gains are desirable in order to have a fast acting DLS. Of course, the values of the gain required would have to be determined by sharpening the estimates. The small value of the coupling constant μ required by the theorem constitutes a problem in practical applications, however. Such small coupling constants would slow down the DLS to an unacceptable degree. Numerical computations of the signal development in the DLS for coupling constants as large as 0.2 have shown that it is not necessary to have a large dominance ratio at the edge of the proportional region; a very strong spectral purification occurs in the final stages of state development, outside the proportional region. This purification is due to the clipping done by the output function, in a manner that is not understood at present. This mechanism is not captured by the present analysis.

Theorem 1 excludes limit cycles from the DLS dynamics. Theorem 2 together with Theorem 3 essentially show that no spurious stable states exist provided that conditions on thresholds, gain, and dimension are met. These conditions are easily satisfied in practice, but the thresholds prescribed are inconsistent with the requirement of gentle nudging of the state into the "correct" gully by the external force, after reset to the origin. Theorem 2 is in agreement with results obtained from the asynchronous discrete DLS model discussed in [9]. For the continuous DLS of practical applications, the large threshold term N^2-4N of (66) must be dropped, and the vector r_a must be taken as specified in (41), with c=N, as determined by (259). Doing this lifts the shield against spurious states, but Theorem 13 assures that the final state is the dominant Hadamard vector, as desired, provided that the coupling constant is as specified and the gain is large enough.

This result can be understood from the features of the energy landscape. The external coupling tilts the energy surface just a little, if the coupling constant μ is small. The tilt also is necessary for the state to move away from the starting point, the origin. For very large values of μ the tilt is so severe as to destroy the stabilty of the Hadamard points. For a range of intermediate values of μ one expects proper behavior of the DLS. It would be valuable to determine this range from theory, since for the practical dimensions N>32 a complete check of associative recall by numerical computation is out of the question because of the computer time required.

Theorem 3 assures the instability of any stationary signal which has components y_a which not all have values near ±1. An example is given by signals with a principal spectrum of even order m. Such a signal has components y_a that are zero, because of the nature of Hadamard vectors. For m=2 the spectrum is conserved in the proportional region, by Theorem 12, and also further out, beyond the proportional region, because the signal trajectory is a straight line through the origin in this case. By Theorems 10 and 11, there is a stationary point with this spectrum, if r=0. By Theorem 3, this stationary point is unstable.

Some of the theorems specify the hyperbolic tangent threshold function, while others use the piecewise linear function. These choices have been made to expedite the proofs. It is

is expected that the results obtained go through for a large class of sigmoid functions.

In regard to the choice between initial value coupling and external coupling, it is noted that from a theoretical point of view the initial value coupling is cleaner, because it does not tilt the energy surface. Then, all Hadamard points are on the same footing in regard to stabilizing forces. Since for zero threshold we have the stable point at the origin (by Theorem 6), the coupling constant μ must be chosen large enough to place the initial signal outside of the crater. Increasing μ means placing the initial signal farther out, and hence has the result of speeding up the DLS action. However, increasing μ also leaves less time for spectral purification, and we may expect a breakdown in associative recall for very large coupling constants. From the practical standpoint, the external coupling is preferred, because it provides for a separate input and output of the DLS. In the presence of noise, the external coupling has the advantage that the BLT output **u** remains standing on the input, whereas the original information **u** is lost soon after reset in the initial value coupling scheme

It may be surprising that so much theory can be developed about a neural net with nonlinear activation; the dynamics in the continuum model is governed by N coupled nonlinear differential equations. Two kinds of nonlinearities are present: the familiar sigmoid nonlinearity in the neuron output function, and the quadratic activation, involving the Hadamard business. The reason for the possibility of extensive theoretical development is, of course, that the stored vectors in the DLS are Hadamard vectors. The properties of these vectors, orthonormality and others, allow deductions and calculations that would not be possible in more general or different settings.

The DLS may be seen as a distributed winner take all circuit. As such, it may have applications beyond the one described here, as second stage of an SRM.

The Hadamard matrices constructed from cyclic S matrices can be generated by shifts of the vector z_{N-1} (see Appendix A). This provides a method of constructing Hadamard vectors in hardware, but for large dimension N the implementation of the vector z_{N-1} becomes cumbersome. Then, one may consider using random bipolar vectors instead of Hadamard vectors. For large dimension, random bipolar vectors have a high probability of being nearly orthogonal, and we may expect most of the theory to go through "on the average". Alternatively, one can start with an index group defined by a structure function which is invariant under permutations, and use this group instead of the Hadamard group. The structure function may be chosen at random, and it would commit the connection tensor S_{abc} in the equations of motion (15). The symmetry of S_{abc} is then a consequence of the invariance of the group structure under index permutations. We see a simple way of generating permutation invariant structure functions by use of $\overline{\text{XOR}}$.

The bipolar vectors that would replace the Hadamard vectors in this scheme need to be computed by the neural net because they must serve as labels for the stored states in the BLT. But if everything goes as expected, such states arise naturally as settled states of the DLS, hooked up according to the group structure. By Hebb learning these states can be impressed, together with the stored states (as outer products) upon the BLT connection matrix B.

APPENDIX A

In this appendix a discussion is given of Hadamard matrices and Hadamard vectors, insofar as needed in this report. *Hadamard vectors* are bipolar vectors that form an orthonormal set The norm of the vectors is not unity, but $\sqrt{N}$. where N is the dimension, here restricted to be a power of 2. A square matrix, the rows of which are Hadamard vectors, is called a *Hadamard matrix*. Hadamard matrices have important application in signal processing involving multiplexing; they are used in spectrometers and imagers to improve the signal to noise ratio. Optics employing this method is called "Hadamard transform optics" [13]. Hadamard matrices are also used in error–correcting codes [13]. They are used here simply because in our SRM there is a need for orthogonal labels that are bipolar vectors.

The Hadamard vectors are denoted by $\mathbf{h}_\alpha$, α=1 to N. Labeling the vector components by the index b=1 to N, the Hadamard matrix with rows $\mathbf{h}_\alpha$ has the elements $h_{\alpha b}$. The orthonormality of the Hadamard vectors is expressed by

$$h^\alpha{}_a h_\alpha{}^b = N\delta_a{}^b , \tag{A1}$$

where $\delta_a{}^b$ is the Kronecker. From (A1) and the linear independence of the Hadamard vectors a second set of orthonormality conditions can be derived:

$$h^\alpha{}_a h_\beta{}^a = N\delta_\beta{}^\alpha . \tag{A2}$$

In this report we restrict the Hadamard matrices to be symmetric, and to have solely + in their first row. For dimensions N that are powers of 2 there are the *Sylvester–type* Hadamard matrices [13], which are defined recusively by the scheme

$$H_{2N} = \begin{vmatrix} H_N & H_N \\ H_N & -H_N \end{vmatrix} , H_2 = \begin{vmatrix} + & + \\ + & - \end{vmatrix} . \tag{A3}$$

For example, The Sylvester–type Hadamard matrix of dimension 16 is shown below

```
+ + + + + + + + + + + + + + + +
+ − + − + − + − + − + − + − + −
+ + − − + + − − + + − − + + − −
+ − − + + − − + + − − + + − − +
+ + + + − − − − + + + + − − − −
+ − + − − + − + + − + − − + − +
+ + − − − − + + + + − − − − + +
+ − − + − + + − + − − + − + + −
+ + + + + + + + − − − − − − − −
+ − + − + − + − − + − + − + − +
+ + − − + + − − − − + + − − + +
+ − − + + − − + − + + − − + + −
+ + + + − − − − − − − − + + + +
+ − + − − + − + − + − + + − + −
+ + − − − − + + − − + + + + − −
+ − − + − + + − − + + − + − − +
```

There is a second type of Hadamard matrix. These matrices are constructed from so–called cyclic S matrices [13], and they exist only for certain dimensions N, including all powers of 2. The construction of a matrix of this type starts with choosing the first Hadamard vectors to have all components +. The remaining N–1 Hadamard vectors are found by taking their first component as +, and the remaining N–1 components as left shifts of a special N–1 dimensional bipolar vector $\mathbf{z}_{N-1}$, the construction of which is discussed by Harwit and Sloane [13]. They also show a list of such vectors for several small values of N. Examples taken from [13] are:

N–1	$\mathbf{z}_{N-1}$
3	− + −
7	− − − + − + +
11	− − + − − − + + + − +
15	+ + + − + + − − + − + − − − −

The Hadamard matrix constructed from $\mathbf{z}$ for N=16 is

$$
\begin{array}{cccccccccccccccc}
+ & + & + & + & + & + & + & + & + & + & + & + & + & + & + & + \\
+ & + & + & + & - & + & + & - & - & + & - & + & - & - & - & - \\
+ & + & + & - & + & + & - & - & + & - & + & - & - & - & - & + \\
+ & + & - & + & + & - & - & + & - & + & - & - & - & - & + & + \\
+ & - & + & + & - & - & + & - & + & - & - & - & - & + & + & + \\
+ & + & + & - & - & + & - & + & - & - & - & - & + & + & + & - \\
+ & + & - & - & + & - & + & - & - & - & - & + & + & + & - & + \\
+ & - & - & + & - & + & - & - & - & - & + & + & + & - & + & + \\
+ & - & + & - & + & - & - & - & - & + & + & + & - & + & + & - \\
+ & + & - & + & - & - & - & - & + & + & + & - & + & + & - & - \\
+ & - & + & - & - & - & - & + & + & + & - & + & + & - & - & + \\
+ & + & - & - & - & - & + & + & + & - & + & + & - & - & + & - \\
+ & - & - & - & - & + & + & + & - & + & + & - & - & + & - & + \\
+ & - & - & - & + & + & + & - & + & + & - & - & + & - & + & - \\
+ & - & - & + & + & + & - & + & + & - & - & + & - & + & - & - \\
+ & - & + & + & + & - & + & + & - & - & + & - & + & - & - & -
\end{array}
$$

In this report we use the Hadamard matrices constructed from cyclic S matrices. The dimension N is restricted to a power of 2. For these Hadamard vectors $\mathbf{h}_\alpha$ we have

$$\sum_a h_{\alpha a} = N\delta_{a1} \; ; \tag{A4}$$

by the symmetry of Hadamard matrices used here, this may also be written as

$$\sum_\alpha h_{\alpha a} = N\delta_{\alpha 1} \; . \tag{A5}$$

We further have

$$h_{1a}=1, \forall a \text{ , and } h_{\alpha 1}=1, \forall \alpha \; . \tag{A6}$$

The Hadamard transform p_α of a vector p_a is given by

$$p_\alpha = (1/\sqrt{N}) h_\alpha{}^a p_a \,. \tag{A7}$$

The same transformation is used for contravariant vectors:

$$p^\alpha = (1/\sqrt{N}) h^\alpha{}_a p^a \,. \tag{A8}$$

Indices are raised and lowered with the Kronecker δ as metric tensor. The Hadamard transform is an orthogonal transformation, i.e., it preserves scalar products. As a result, one has

$$y_\alpha y^\alpha = y_a y^a \,. \tag{A9}$$

The inverse Hadamard transforms are:

$$p_a = (1/\sqrt{N}) h^\alpha{}_a p_\alpha \,, \tag{A10}$$

$$p^a = (1/\sqrt{N}) h_\alpha{}^a p^\alpha \,. \tag{A11}$$

For any vector **p**, the p_a may be seen as the components in the neuron frame, whereas the p_α are the components of the vector in the Hadamard frame.

In the sequel we need the group property of the Hadamard vectors. This property is discussed in detail in [9], and we only give here the results without proof. We have

Theorem A1: For dimensions that are a power of 2, the Hadamard vectors constructed from cyclic S matrices form a group under component wise multiplication.

This means that

$$h_{\alpha a} h_{\beta a} = h_{\gamma a} \quad , \forall\, a \,, \tag{A12}$$

where

$$\gamma = f(\alpha,\beta) \,; \tag{A13}$$

f is called the *structure function* of the group. The component wise multiplication is a logical $\overline{\text{XOR}}$, so that (A13) may also be expresed as

$$\mathbf{h}_\alpha \,\overline{\text{XOR}}\, \mathbf{h}_\beta = \mathbf{h}_\gamma \,. \tag{A14}$$

The stucture function can be determined from the Hadamard vectors. For example, we have, for N=16, f(2,3)=6, and f(2,4)=10, as can be seen from the Hadamard matrix on page 53. The first Hadamard vector, $\mathbf{h}_1$, has all components + and therefore acts as the identity. It is easy to see that every Hadamard vector is its own inverse,

$$h_{\alpha a} h_{\alpha a} = h_{1a} \,, \forall\alpha, \forall a \,. \tag{A15}$$

For the structure function, (A15) implies

$$f(\alpha,\alpha) = 1 \,, \quad \forall\, \alpha \,. \tag{A16}$$

Moreover, $f(\alpha,\beta)=1 \implies \alpha=\beta$, (A17)

as can be seen as follows. $h_{\alpha a}h_{\beta a}=h_{1a}$,$\forall$ a implies

$$h_{\alpha a}h_{\beta a}h_{\beta a}=h_{1a}h_{\beta a}\,. \tag{A18}$$

Because of the symmetry of the Hadamard matrices used, the roles of indices α and a may be reversed, so that we also have

$$h_{\alpha a}h_{\alpha b}=h_{\alpha c}\,, \tag{A19}$$

where

$$c=f(a,b)\,. \tag{A20}$$

The function f is symmetric,

$$f(a,b)=f(b,a), \tag{A21}$$

and (A20) is invariant under cyclic permutation of indices,

so that (A20) implies

$$b=f(c,a)\,. \tag{A22}$$

APPENDIX B

If the connection tensor is fully symmetric, the velocity vector $\dot{\mathbf{v}}$ given by (15) is curlfree in the space of signals **y**, but not in activation space. This is shown as follows.

Considering $\dot{v}_a$ as a function of **y**, we have from (15)

$$\partial\dot{v}_a/\partial y_d - \partial\dot{v}_d/\partial y_a =$$

$$-\delta_{da}/s'(v_d)+\delta_{ad}/s'(v_a)+(S_{adb}+S_{abd})y^b-(S_{dab}+S_{dba})y^b = 0\,, \tag{B1}$$

by the symmetry of the connection tensor, and the property of the Kronecker δ. However, if $\dot{v}_a$ is considered to be a function of **v** we have

$$\partial\dot{v}_a/\partial v_d-\partial\dot{v}_d/\partial v_a=-\delta_{da}+\delta_{ad}+(S_{adb}+S_{abd})y^b s'(v_d)-(S_{dab}+S_{dba})y^b s'(v_a), \tag{B2}$$

and this does not vanish if $s'(v_d)\neq s'(v_a)$. This can be remedied by multiplying $\dot{v}_a$ by $s'(v_a)$; one has

$$\partial(\dot{v}_a s'(v_a))/\partial v_d-\partial(\dot{v}_d s'(v_d))/\partial v_a=\dot{v}_a\delta_{da}s''(v_a)-\dot{v}_d\delta_{ad}s''(v_d)-s'(v_a)\delta_{da}+s'(v_d)\delta_{ad}$$

$$+(S_{adb}+S_{abd})y^b s'(v_a)s'(v_d)-(S_{dab}+S_{dba})y^b s'(v_d)s'(v_a)\,. \tag{B3}$$

The first two terms cancel each other because of the Kronecker delta. Similarly, the third and fourth terms cancel. The S terms now cancel because each has the same factor $s'(v_a)s'(v_d)$, and, of course, the symmetry is needed as well. It follows that the vector field $\dot{v}_a s'(v_a)$ is curlfree in activation space. If we follow this up and take the path integral of

$\dot{v}_a s'(v_a)$ in activation space we get the same Liapunov function as before:

$$-\sum_a \int_0^{v} \dot{v}_a s'(v_a)\, dv_a = -\int_0^{y} \dot{v}_a\, dy^a \,, \tag{B4}$$

by a change of integration variables. On the right the summation convention is used; we could not do that on the left because there are too many indices a.

The discussion clarifies the somewhat unnatural construction $-\int_0^{y} \dot{v}_a dy^a$ for the Liapunov function (20), which is a path integral of the activation velocity vector in signal space.

APPENDIX C

In this appendix we derive some properties and consequences of the neuron output function given by (18),

$$y=s(v) = \tanh(gv) \,, \tag{C1}$$

where g is the gain at zero. The inverse mapping, from y to v, is

$$v = \frac{1}{2g} \ln \frac{1+y}{1-y} \,. \tag{C2}$$

One has

$$s'(v) = g\, \mathrm{sech}^2(gv) = g(1-s^2) \,, \tag{C3}$$

The critical activation v^* was introduced in the sequel for estimation purposes; its value is chosen such that

$$s(v^*) = 1-\epsilon \,, \tag{C4}$$

where ϵ is a small positive number. For the hyperbolic tangent (C1) one has from (C2) for small ϵ

$$v^* \simeq \frac{1}{2g} \ln (2/\epsilon) \,, \tag{C5}$$

$$v^* < \frac{1}{2g} \ln(2/\epsilon) \,, \tag{C6}$$

and

$$s'(v^*) = g\epsilon(2-\epsilon) \simeq 2g\epsilon \,. \tag{C7}$$

In the proportional region, the signal is proportional to the activation, either approximately or exactly. For the hyperbolic tangent output function, we need to know the accuracy of the linear approximation. This accuracy may be calculated as the ratio of the cubic term to the linear term in the power series expansion of the hyperbolic tangent:

$$y=\tanh(gv)=gv-g^3v^3/3+ \,.. \tag{C8}$$

The ratio of the cubic term to the linear term is

$$g^2v^2/3 \simeq y^2/3. \qquad \text{(C9)}$$

From (24) one has

$$\varphi(v) = vs - \int_0^v s(\xi)\, d\xi = v \tanh(gv) - \int_0^v \tanh(g\xi)\, d\xi =$$
$$= v \tanh(gv) - (1/g) \ln \cosh(gv) . \qquad \text{(C10)}$$

If the energy E of (22) is considered a function in signal space, φ needs to be expressed in terms of $\mathbf{y}$. This can be done by using (C2) in (C10):

$$\varphi = \frac{y}{2g} \ln \frac{1+y}{1-y} + \frac{1}{2g} \ln (1/\cosh^2(gv)) =$$
$$= -\frac{1}{2g} \{\ln \frac{1+y}{1-y} + \ln (1 - \tanh^2(gv))\} = \frac{1}{2g} \{\ln \frac{1+y}{1-y} + \ln (1-y^2)\} =$$
$$= \frac{1}{2g} \{(1+y) \ln(1+y) + (1-y) \ln(1-y)\}. \qquad \text{(C11)}$$

Hence, for the function Ω in the energy (22) we have

$$\Omega = \frac{1}{2g} \sum_a \{(1+y_a) \ln(1+y_a) + (1-y_a) \ln(1-y_a)\} . \qquad \text{(C12)}$$

In the sequel we need a polynomial approximation of Ω near the origin. Such an approximation good to powers y_a^3 is

$$\Omega \simeq R^2/(2g) , \qquad \text{(C13)}$$

where

$$R^2 = y^a y_a . \qquad \text{(C14)}$$

For the piecewise linear output function

$$\begin{aligned} s(v) &= -1 \text{ if } v < -1/g , \\ &= gv \text{ if } |v| \leq 1/g , \\ &= 1 \text{ if } v > 1/g , \end{aligned} \qquad \text{(C15)}$$

the proportional region is given by

$$|v| < 1/g . \qquad \text{(C16)}$$

In the proportional region we have for the function Ω

$$\Omega = R^2/(2g) ; \qquad \text{(C17)}$$

this can be verified by taking partial derivatives:

$$\partial\Omega/\partial y^a = y_a/g = v_a , \qquad \text{(C18)}$$

as required.

APPENDIX D

For convenience we show here the proof of Lemma 1, which is just Theorem 2 of [9]. The proof is formally altered at some places because of a different factor in the Hadamard transform.

It should be noted that in this appendix the signal **y** is restricted to be bipolar, i.e., **y** is a corner point of the solid hypercube J_N.

The material considered involves the vector Q_a defined by (54)

$$Q_a = N \sum_{\alpha} h_{\alpha a} y_\alpha^2 . \tag{D1}$$

where

$$y_\alpha = (1/\sqrt{N})\, h_\alpha^{\ a} y_a \tag{D2}$$

is the Hadamard transform of the signals y_a.

Before stating and proving Lemma 1, we need some preparation. Define the index set A:

$$A=\{\alpha | y_\alpha \neq 0\} , \tag{D3}$$

and the disjoint pieces

$$A_a^+=\{\alpha | \alpha \in A \ \& \ h_{\alpha a}=+\} , \tag{D4}$$

and

$$A_a^-=\{\alpha | \alpha \in A \ \& \ h_{\alpha a}=-\} . \tag{D5}$$

Using this decomposition, (D1) may be written

$$Q_a/N= \sum_{\alpha \in A_a^+} y_\alpha^2 - \sum_{\alpha \in A_a^-} y_\alpha^2 . \tag{D6}$$

From (A9) and the fact that the components y_a are ± 1, one has

$$\sum_{\alpha} y_\alpha^2 = N ; \tag{D7}$$

this may be written as

$$N= \sum_{\alpha \in A_a^+} y_\alpha^2 + \sum_{\alpha \in A_a^-} y_\alpha^2 . \tag{D8}$$

(D6) and (D8) give

$$2 \sum_{\alpha \in A_a^+} y_\alpha^2 = N + Q_a/N , \tag{D9}$$

and

$$2 \sum_{\alpha \in A_a^-} y_\alpha^2 = N - Q_a/N ; \tag{D10}$$

since the left hand sides are nonnegative, this implies

$$-N^2 \le Q_a \le N^2 . \tag{D11}$$

We show

Lemma D1: $\mathbf{y}\epsilon I_N$, $a\neq 1$, $Q_a=-N^2$ <===> $h_{\alpha a}=-$, $\forall\ \alpha\epsilon A$.

Proof: To show <=== , we note that $h_{\alpha a}=-$, $\forall\ \alpha\epsilon A$, implies that $a\neq 1$, by (A6), and

$$Q_a = N \sum_\alpha h_{\alpha a} y_\alpha^{\,2} = -N \sum_\alpha y_\alpha^{\,2} = -N^2 , \tag{D12}$$

by (D7).
To show ==> , we note that $Q_a=-N^2$ and (D9) imply that the set A_a^+ is empty; i.e., $h_{\alpha a}=-$, $\forall\ \alpha\epsilon A$. ⌋

After this preparation we proceed with

Lemma 1: $\mathbf{y}\epsilon I_N$, $Q_b=-N^2$, $\forall b$ such that $y_b=-$ <===> $\mathbf{y}$ is Hadamard $\neq \mathbf{h}_1$.

Proof: Define the index set

$$B=\{b \mid y_b=-\} . \tag{D13}$$

Then the Hadamard transform (D2) may be written as

$$y_\alpha = (1/\sqrt{N}) \{-\sum_{b\epsilon B} h_\alpha^{\,b} + \sum_{b\notin B} h_\alpha^{\,b}\} . \tag{D14}$$

The property (A5) of the Hadamard vectors used here implies that

$$\sum_{b\notin B} h_\alpha^{\,b} = N\,\delta_{\alpha 1} - \sum_{b\epsilon B} h_\alpha^{\,b} , \tag{D15}$$

so that (D14) may be written

$$\sqrt{N}\, y_\alpha = N\delta_{\alpha 1} - 2\sum_{b\epsilon B} h_\alpha^{\,b} . \tag{D16}$$

Suppose $Q_b=-N^2$, $\forall\ b\epsilon B$. Then, Lemma D1 and (D16) give

$$\alpha \in A\ ,\ \sqrt{N}\, y_\alpha = N\delta_{\alpha 1} + 2(N-W)/2 = N\delta_{\alpha 1} + N - W , \tag{D17}$$

where

$$W = \sum_a y_a . \tag{D18}$$

and $(N-W)/2$ is the cardinality of the set B. Hence we have

Lemma D2: $Q_b = -N^2$, $\forall\ b\epsilon B$ ===> $\sqrt{N}\, y_\alpha = N\delta_{\alpha 1} + N - W$, $\forall \alpha\epsilon A$.

There appear to be two cases, Case 1: $\alpha=1 \in A$, and Case 2: $\alpha=1 \notin A$.

Case 1: $\alpha=1 \in A$. Calculating y_α of Lemma D2, we have

$$\sqrt{N}\, y_1 = 2N - W \,. \tag{D19}$$

Also, from (D2), (D18), and the fact that $\mathbf{h}_1$ has all + components we have

$$\sqrt{N}\, y_1 = W \,. \tag{D20}$$

From (D19) and (D20) it follows that

$$W = N \,, \tag{D21}$$

so that $\mathbf{y} = \mathbf{h}_1$, the all positive Hadamard vector. However, for $\mathbf{y} = \mathbf{h}_1$ (D1) gives $Q_b = N^2$, in contradiction with $Q_b = -N^2$ asssumed in Lemma D2. Hence, Case 1 is not possible within the premises of Lemma D2.

Case 2: $\alpha=1 \notin A$. From (D2) for $\alpha=1$ and (D18) it follows that W=0. Hence, for $\alpha \neq 1 \in A$ the y_α of Lemma D2 is just $\sqrt{N}$. With (A9) it follows that

$$N = \sum_\alpha y_\alpha^{\,2} = \sum_{\alpha \in A} y_\alpha^{\,2} = r\, N \,, \tag{D22}$$

where r is the cardinality of set A. Since (D12) implies r=1, the set A contains only a single element, say, γ. It follows that $\mathbf{y} = \mathbf{h}_\gamma$, and we may conclude

$$Q_b = -N^2 \,,\ \forall\ a \in B \ \Longrightarrow\ \mathbf{y} \text{ is Hadamard} \neq \mathbf{h}_1 \,, \tag{D23}$$

which is the forward part of Lemma 2.

The converse is also true, since for $\mathbf{y}=\mathbf{h}_\gamma$, $\gamma \neq 1$, we have $y_\alpha=0$, $\forall\ \alpha \neq \gamma$, and $y_\gamma=\sqrt{N}$, so that (D1) gives

$$Q_b = N^2 h_{\gamma b} \,. \tag{D24}$$

For index b such that $y_b = h_{\gamma b} = -$ it follows that $Q_b = -N^2$. ⌋

REFERENCES

[1] J. J. Hopfield, "Neural Networks and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci. USA **79**, 2554 (1982)
[2] S. Grossberg, Biological Cybernatics **23**, 187, (1976)
[3] D. J. Amit, H. Gutfreund, and H. Sompolinsky, "Information storage in neural networks with low levels of activity", Phys. Rev. **A35**, 2293 (1987)
[4] L. Personnaz, I. Guyon, and G. Dreyfus, "Collective computational properties of neural networks: New learning mechanisms", Phys. Rev. **A34**, 4217 (1986)
[5] I. Kanter and H. Sompolinski: "Associative recall of memory without errors", Phys. Rev. **A35**, 380 (1987)
[6] C. M. Bachman, L. N. Cooper, A. Dembo, and O. Zetouni, Proc. Natl. Acad. Sci. USA **21**, 7529 (1987)
[7] B. Kosko, "Bidirectional Associative Memories", IEEE Trans. Systems, Man,

Cybernatics, **SMC–18**, 49 (1988)
[8] H. G. Loos, in NEURAL INFORMATION SYSTEMS, Ed. Dana Z. Anderson, p. 495, American Institute of Physics, New York, 1988
[9] H. G. Loos, "Quadratic Hadamard Memories", Technical Report #1, ARPA Order No. 6429, Contract DAAH01–88–C–0887, Dec. 1989
[10] J. A. Schouten, RICCI CALCULUS, SPRINGER, NEW YORK, 1954
[11] H. H. Chen, Y. C. Lee, G. Z. Sun, and H. Y. Lee, "High Order Correlation Model for Associative Memory", NEURAL NETWORKS FOR COMPUTING, AIP Conference Proceedings # 151, p.86, (1986)
[12] D. J. Volper and S. E. Hampson, "Quadratic Function Nodes: Use, Structure, and Training", Neural Networks **3**, 93 (1990)
[13] M. Harwit and N. J. A. Sloane, HADAMARD TRANSFORM OPTICS, Academic Press, New York, 1979

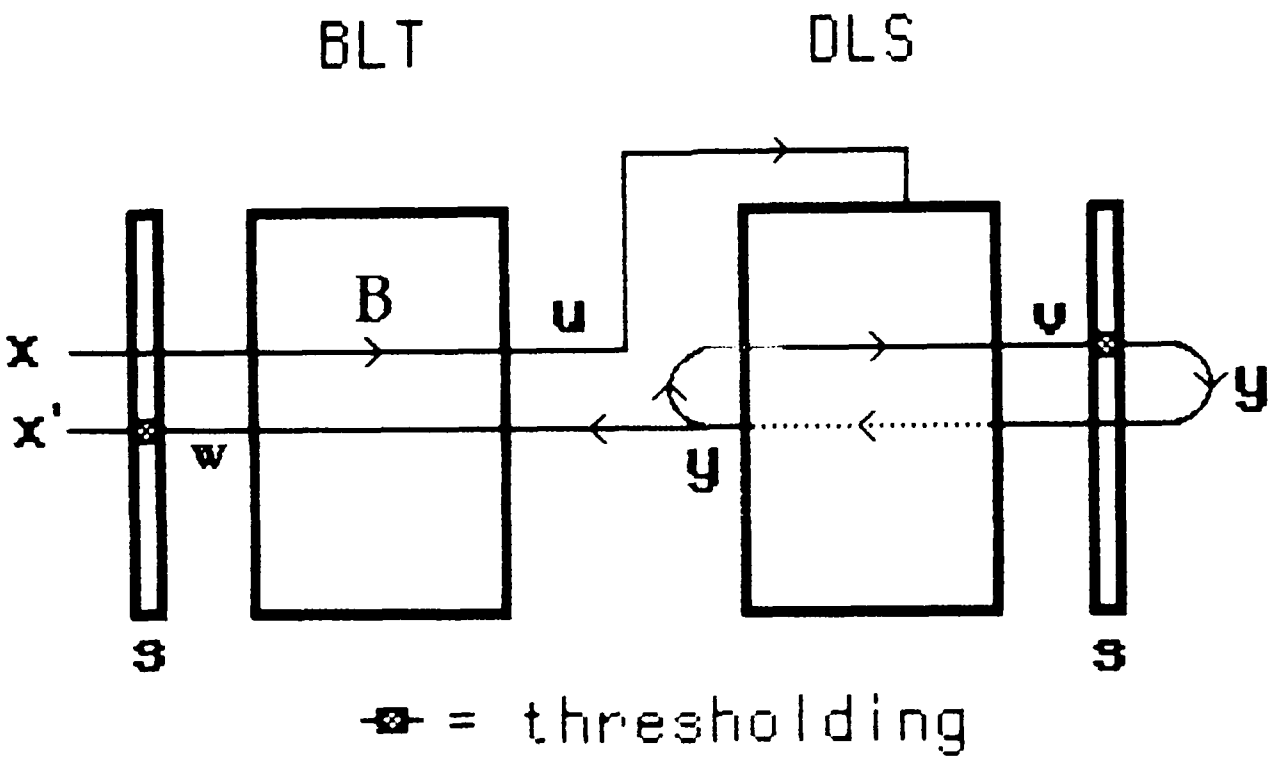

Fig. 1 Selective Reflexive Memory (SRM), consisting of a Bidirectional Linear Transformer (BLT) and a Dominant Label Selector (DLS). The input vector **x** is bipolar. The BLT performs a linear transformation **u**=B**x**. The BLT output **u** is presented to the DLS, which selects from **u** the Hadamard vector **y**=$\mathbf{h}_\beta$ that occurs with the largest coefficient in the Hadamard expansion of **u**. The selected Hadamard vector $\mathbf{h}_\beta$ is returned to the BLT, and is processed in the BLT backstroke to give **w**=$\mathbf{h}_\beta$B. **w** is thresholded to give **x'**. Things can be arranged such that **x'** is the stored vector nearest the input **x**.